Adaptation and attention in higher visual perception

D I S S E R T A T I O N

for the award of the degree
"Doctor rerum naturalium"
Division of Mathematics and Natural Science
of the Georg-August-Universität Göttingen

submitted by

Daniel Kaping
from Berlin

Göttingen 2009

Doctoral thesis committee:

Prof. Dr. Stefan Treue (Advisor, First Referee)
Abt. Kognitive Neurowissenschaften
Deutsches Primatenzentrum (DPZ)
Kellnerweg 4
37077 Göttingen

Dr. Alexander Gail (Second Referee)
Sensorimotor Group, BCCN
Deutsches Primatenzentrum (DPZ)
Kellnerweg 4
37077 Göttingen

Dr. Peter Dechent
MR-Forschung in der Neurologie und Psychiatrie
Universitätsmedizin Göttingen
Georg-August-Universität
Robert-Koch-Str. 40
37075 Göttingen

External thesis advisor:

Prof. Dr. Julia Fischer
Abt. Kognitive Ethologie
Deutsches Primatenzentrum (DPZ)
Kellnerweg 4
37077 Göttingen

Prof. Dr. Uwe Mattler
Abteilung für Experimentelle Psychologie
Georg-Elias-Müller-Institut für Psychologie
Georg-August-Universität
Goßlerstr. 14
37073Göttingen

Prof. Dr. Fred Wolf
Theoretical Neurophysics, BCCN
Max Planck Institute for Dynamics and Self-Organization
Bunsenstrasse 10
37073 Göttingen

Date of submission of the thesis: 31 December, 2009

Date of disputation: 17 February, 2010

I hereby declare that this thesis has been written independently and with no other sources and aids than quoted.

Göttingen, 31 December, 2009 Daniel Kaping

Acknowledgments

I would like to thank Stefan Treue for giving me the possibility to work in his laboratory and to study under his supervision. I am very grateful for the help, guidance and numerous opportunities offered during the course of the years. Alexander Gail and Peter Dechent, both members of my PhD committee, have always offered important advice and constructive criticism. I also thank Julia Fischer, Uwe Mattler and Fred Wolf for their kind support in evaluating this thesis.

For the electrophysiological part of this work, I would like to thank Leonore Burchardt, Sina Plümer and Dirk Prüsse for expert support regarding all questions of animal care. A special note of gratitude goes to Sonia Baloni for frequent help with the electrophysiological recordings and taking care and charge of Nico and Wallace.

I am also very happy to thank Sabine Stuber and Beatrix Glaser for all administrative work; further Ralf Brockhausen and Kevin Windolph for their computer and technical support.

I am grateful to Laura Busse and Steffen Katzner for all their excelnet advise and help; toTzvetomir Tzvetanov and Carmen Morawetz for productive discussions.

I would like to thank Stephanie, Anja, Lu, Valeska, Vladislav, Shubo, Katharina, Pinar, Robert, Florian, Thilo who not only provided intellectual and emotional support but made the laboratory a fun place to be.

Last but not least I thank my mother, Eva and my sister, Daniela for their continuous motivation, encouragement and support.

And, thank you Johanna for your loving, supportive spirit, your voice of reason, you are my life raft.

Contents

Chapter I

Introduction

Seeing, at its simplest, is merely the registering of light and some reaction to it. Primate visual perception is not only a passive, feedforward absorption of information of the surrounding environment. While simple light sensitive creatures show purely stimulus driven light avoidance / attraction responses, our own visual system consists not only of “low-level” vision but of more complex mechanisms operating on the “low-level” output. It is the interpretation of what we see in the light of knowledge and experience about the world. Vision is therefore also influenced by intention, context and memory. These do not make their contribution late within the visual processing chain but rather affect all cortical processing of visual input.

This thesis identifies two mechanisms, visual adaptation and visual attention, to shape sensory information in the visual system giving rise to conscious perception. “High-level”, conscious perception describe in part our ability to recognize objects such as faces and navigate / orientate within the world. Out of the vast amount of “low-level” information captured by our eyes only a small, selected fraction reaches consciousness. Despite the astonishing large amount of cortex dedicated to visual perception (roughly 50% of the macaque and between 20 - 30% of the human cortex are dedicated to vision; Orban, VanEssen and Vanduffel, 2004) the possible changes along many dimensions of a given stimulus (e.g positioning, orientation, motion, lighting conditions etc.) require effective recalibration (adaptation) and filter (attention) mechanisms enhancing the behaviorally most important stimulus or stimulus attributes.

To study the adaptive influences on visual perception I have made use of psychophysical methods and functional magnetic resonance imaging (fMRI). Recent findings show that individual environmental scenes can be classified by their underlying statistical properties. Natural scenes differ from scenes containing man-made

environments along their frequency profile. Can this classification of complex scenes be adaptively influenced by “low-level” statistical properties to which the observer is exposed to? Two psychophysical studies presented in this thesis (chapter 2) suggest that the classification of man-made and natural images can routinely be influenced by the statistical scene properties of the individual’s environment. The prolonged exposure to “low-level” stimuli is known to produce perceptual aftereffects; surprisingly adaptation to complex face stimuli resemble “low-level” adaptational adjustments. Does this adaptive recalibration of “high-level” face perception normalize to common properties in the environment allowing for a common state and shared visual experiences? The fMRI face distortion aftereffect study presented in the later part of chapter 2 illustrates the effect of norm-based face adaptation, providing evidence of neural responses coupled to illusory post-adaptive face percepts .

Attentional effects on the processing of sensory information were explored within a model system, the highly developed ability of primates to process visual motion. I have performed extracellular recordings in the motion sensitive medial superior temporal area (MST) of two awake behaving macaque monkeys. MST neurons receive their primary input from motion sensitive middle temporal area (MT). While MST neurons respond to linear motion, tuning to complex spiral motion stimuli is more pronounced. Little is known about whether and how MST responses change with attention. Here, the focus will mainly be on spatial and feature-based attention; “top-down” mechanisms known to modulate the processing of sensory information.

As the studies presented in this thesis are based upon mechanisms related to visual perception a short overview of the primate visual system will be provided. The main part of this work will be divided into separate chapters: Adaptation and Attention; each of these subsections consisting of original research articles and manuscripts. Brief descriptions of visual adaptation (chapter 2) and attention (with emphasize on visual motion processing in area MST; chapter 3) will be given. The experiments main objectives and major findings will briefly be introduced in a preceding section of each manuscript.

I.I The primate visual system

Even with the human cortex surface area spanning over 10 times of that of the macaque cortex surface area (VanEssen, Harwell, Hanlon and Dickson, 2005) several cortical regions have been identified to be homologous between the two species. The largest coherence has been recognized within cortical areas dedicated to the processing of vision. Information from the retina travels via a part of the thalamus called the lateral geniculate nucleus (LGN) to the primary visual cortex, also known as visual area one (V1). Points that are next to each other on the retina connect to cells next to each other in V1. Cells in V1 also connect back to the LGN, and this feedback neural traffic is characteristic of the entire visual system. The primary visual cortex V1 is only the first of several visual areas in the occipital lobe. In both macaque and human V1 is the single largest area dedicated to the processing of vision (10% macaque and 3% human; VanEssen, 2005). V1 cell activation is tightly coupled to specific stimulus properties such as edges and borders; the gradient illumination of a bar type stimulus. The V1 sensory inputs not only allow for selective processing of orientation and direction but also code information about stimulus color. V1 is only the very first step in the hierarchical organized processing of vision. Two main visual pathways leave area V1: (i) the ventral pathway conveying information to the temporal lobe (V1, V2, V4, TEO, IT), specialized for the processing color, shape and object identity, and (ii) the dorsal pathway projecting to the parietal cortex (V1, V2, V3, MT, MST, LIP), processing information about motion, spatial relations and depth.

Within the hierarchy of cortical visual processing the information transformation from a simple bar / line stimulus to increasingly more complex visual objects in the environment is based upon receptive fields (RFs). RFs are single cell spatial restricted response regions (relative to the fovea), progressively increasing in size from one visual area to the next. Integrating more visual information as the complexity of the preferentially coded stimulus attributes of the RFs change along the cortical visual processing hierarchy; from well understood “low-level”, oriented line, V1 responses to: (i) “high-level” increasingly more complex motion patterns within MT / MST along the dorsal pathway; (ii) “high-level” single face selective neurons within the temporal lobe (ventral pathway).

Chapter II

Adaptation

"This must also happen in the organ wherein sense-perception takes place, since sense-perception , as realized in actual perceiving, is a mode of qualitative change. (...) after having looked at the sun or some other brilliant object, we close the eyes, then if we watch carefully, it appears in a right line with the direct vision, at first at its own colour; then it changes to crimson, next to purple, until it becomes black and disappears. And also when persons turn away from looking at objects in motion (...) they find the visual stimulations still present themselves, for the things really at rest are then seen moving: (...) sensory organs are acutely sensitive to even a slight qualitative difference (...) and that sense-perception is quick to respond to it; and further that the organ which perceives is not only affected by its object, but also reacts to it."

Aristotle

(On Dreams)

Visual adaptation is an unconscious process of adjustment of the visual system to its environment. A dynamic attempt to preserve sensitivity to potential changes. Our visual system could not possibly be so sensitive to small increments and decrements of a stimulus signal if the whole range of possible changes had to be encoded. Barlow (1990) describes the sensory signal "*the spike train*" to be "*a somewhat crude method of signaling a metric quantity; the number of reliably distinguishable levels of activity in a small time interval is very limited, so the distinguishable steps... would be very large without the adaptive mechanism (...)*"

These adaptive response changes were believed to be some form of fatigue in a cells response to the repeatedly exposure to the same stimulus (Sekuler and Pantle, 1967;

Vautin and Berkley, 1977). Carandini (2000) pointed out that there has to be more than neural fatigue to adaptation; as fatigue should affect the responses to all stimuli equally. Instead the largest suppression can be observed when adaptation and test stimulus match the preferred stimulus of a given cell (Movshon and Lennie, 1979), while adapting to the anti-preferred enhances response to the preferred stimulus (Petersen, Baker and Allman, 1985). These adjustments preserving sensitivity to small variations in the visual environment and removing redundancies take place at the expense of accurate representation of the environment.

Adaptation induced perceptual aftereffects can make us aware of the fact that perception is not a window onto reality. They not only occur early at the receptor level (light / dark adjustments) but also at "higher stages" of the visual system adjusting complex image properties sometimes triggering illusory motion / figural aftereffects. These visual illusions expose the adaptive adjustments made by the visual system to prolonged viewing of a given stimulus set. The visual inaccuracies resulting from the dynamic response range of the "newly" adapted visual system have been studied in an attempt to understand how the brain processes certain visual information (e.g. orientation selectivity (Graham, 1972), direction selectivity (Tootell et al. 1995), color opponency (Webster and Mollon, 1994) and figural aftereffects (Webster & McLin, 1999; Rhodes, Jeffery Watson, Clifford & Nakayama, 2003; Watson & Clifford, 2003)). Perceptual aftereffects abide by time-courses of logarithmic build-up and exponential decay (Rhodes, Jeffrey, Clifford & Leopold, 2007; Leopold, Rhodes, Müller & Jeffrey, 2005). Typically these aftereffects bias perception towards the opposite of the adapting stimulus resulting in a recalibration of the visual system establishing a new neutral point according to the average of the prevailing stimulus (Clifford, Webster, Stanley, Stocker, Kohn, Sharpee and Schwartz, 2007).

Recent studies have proposed an adaptational recalibration adjustment to encode stimuli not in terms of their absolute structure but as a deviation from a set norm (Webster, Werner & Field, 2005). If perceptual adjustments center around a well established norm to what extend are these adjustments molded around the same or different environments we are exposed to? The following manuscripts test for adaptational adjustments to statistical properties within natural / man-made environmental scenes and a norm-based face aftereffect within human observers.

Adaptation - original articles and manuscripts

- Kaping D, Tzvetanov T and Treue S (2007). Adaptation to statistical properties of visual scenes biases rapid categorization. Visual Cognition; 15: 12-19

  Author contribution: DK. and TT designed and performed the experiment; DK wrote the main paper, and TT wrote the Methods section. ST edited the manuscript; all authors discussed the results and commented on the manuscript at all stages.

- Kaping D and Treue S. Adaptation to image statistics decreases sensitivity to the prevailing scene. Prepared for submission

  Author contribution: DK designed and performed the experiment; DK wrote the manuscript and ST edited the manuscript; all authors discussed the results and commented on the manuscript at all stages.

- Kaping D, Morawetz C, Baudewig J, Treue S, Webster MA and Dechent P. The face distortion aftereffect reveal norm-based coding in human face perception. (submitted)

  Author contribution: DK and MW designed the original experiment, MW developed stimuli; CM and JB implemented the fMRI experiment. CM and DK collected and analyzed data. DK and MW wrote the main paper, and CM wrote the Methods section. MW, JB, ST and PD edited the manuscript; all authors discussed the results and commented on the manuscript at all stages.

II.I Adaptation to statistical properties of visual scenes biases rapid categorization

Object and scene recognition display the remarkable ability of the human visual system to recognize complex, continuously changing environments. Despite an extensive amount of information being presented within a given environmental scene, early categorization involving man-made / natural object detection is carried out effortlessly requiring little to no attention. Rapid and parallel categorization of novel scenes and objects is believed to be dependent upon higher-level cortical areas, such as infero-temporal cortex, responding to various categories of objects. Can hierarchical processing of low-level, simple stimulus attributes be a sufficient tool in the processing of everyday visual scenes?

Torralba and Oliva (2003) propose that with respect to natural environments, the power spectrum for scenes containing man-made environments differ along their frequency profile. The power spectrum is the amount of a given 2D spatial frequency for a specific orientation contained in the image. Natural environmental images cover a broad variation in spectral shapes whereas man-made environments mainly differ along horizontal and vertical contours. Irregularities underlying these statistical properties of different environments could require only minimal processing time and may account for rapid scene / object categorization. While untested, this provides the basis for a plausible image recognition mechanism based upon a low-level feedforward process within the early visual system (namely V1 and V2).

Based upon “low-level” features (statistical properties) of different environmental categories, we employed an adaptation paradigm to test the contribution of early process within the visual system. Adaptation to artificial images mimicking the underlying statistical properties of an environmental scene recalibrated the human visual system at a very early stage and alter the perception of a subsequently viewed environment. This suggests that the classification of man-made and natural images can be based upon a feedforward system routinely influenced by “low-level” statistical properties.

VISUAL COGNITION, 2007, 15 (1), 12–19

Psychology Press
Taylor & Francis Group

Adaptation to statistical properties of visual scenes biases rapid categorization

Daniel Kaping, Tzvetomir Tzvetanov and Stefan Treue

Cognitive Neuroscience Laboratory, German Primate Centre, Goettingen, Germany

The initial categorization of complex visual scenes is a very rapid process. Here we find no differences in performance for upright and inverted images arguing for a neural mechanism that can function without involving high-level image orientation dependent identification processes. Using an adaptation paradigm we are able to demonstrate that artificial images composed to mimic the orientation distribution of either natural or man-made scenes systematically shift the judgement of human observers. This suggests a highly efficient feedforward system that makes use of "low-level" image features yet supports the rapid extraction of essential information for the categorization of complex visual scenes.

The human visual system has a remarkable ability to recognize objects, even in the midst of complex, continuously changing environments. This requires the transformation of a point-by-point retinal image into the neuronal representation of an object that is view-invariant, i.e., largely unaffected by changes in position, orientation, distance, or the presence of other visual objects in the vicinity. The recognition and categorization of scenes and objects is believed to be performed in higher level cortical areas such as the inferotemporal cortex (Logothetis & Sheinberg, 1996; Tanaka, 1996) and the medial temporal lobe (Kreiman, Koch, & Fried, 2000).

Despite its inherent difficulty, detection and categorization of objects and scenes is carried out effortlessly (Li, VanRullen, Koch, & Rerona, 2002), remarkably fast (Grill-Spector & Kanwisher, 2005; Potter, 1976), and is robust to manipulations such as image inversion (Rousselet, Mace, & Fabre-Thorpe, 2003). In a series of experiments Thorpe and colleagues (Rousselet, Fabre-Thorpe, & Thorpe, 2002; Thorpe, Fize, & Marlot, 1996; VanRullen &

Please address all correspondence to Stefan Treue, Cognitive Neuroscience Laboratory, German Primate Centre, Kellnerweg 4, 37077 Goettingen, Germany. E-mail: treue@gwdg.de

This research project has been supported by a Marie Curie Early Stage Research Training Fellowship of the European Community's Sixth Framework Programme under the contract number MEST-CT-2004-007825.

http://www.psypress.com/viscog DOI: 10.1080/13506280600856660

Thorpe, 2001) asked human subjects to decide whether an unmasked picture of a scene presented for only 20 ms contained an animal or not. Measuring event related potentials the authors were able to document different frontal activation between the two picture types only 150 ms after stimulus onset, suggesting that this type of categorization is relying on a feedforward mechanism, rather than on a high-level feature detection system located high up in the visual processing hierarchy (Rousselet et al., 2003).

Such findings point to a system that can rely on low-level image analysis for accurate object detection and scene categorization. Several factors can contribute to such a system: It has been pointed out that the general layout of scenes supports scene recognition after only a short glance (Friedman, 1979). A correct category detection permits an overall scene evaluation along more general, superordinate levels allowing the extraction of categorical properties of the depicted scene independent of detailed object recognition (Biederman, 1981; Oliva & Torralba, 2001).

Additionally, simple hierarchical processing can build upon easily extractable statistical image information (Oliva & Schyns, 1997), such as the spatial frequency composition of an image extracted through image decomposition via Fourier transformation, and the use of the orientation-selective neurons in early visual cortex. This would provide a plausible mechanism for the rapid categorization process.

For such an approach to work, scenes that are to be distinguished should differ in their respective Fourier spectra and these differences need to be large enough to enable reliable scene categorization. Indeed, Torralba and Oliva (2003) showed that the power spectrum of natural environments differ from man-made environments (Figure 1), particularly because of the

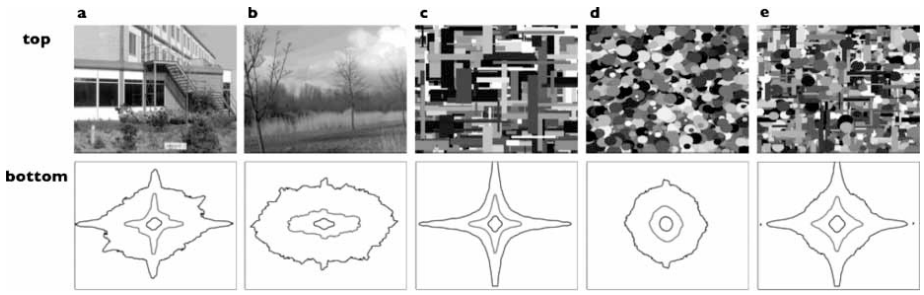


Figure 1. Examples of the images (top row) used in this study with their corresponding power spectrum (bottom row, see also Torralba & Oliva, 2003). The contour plots represent 70% (outer line), 80% (middle line), and 90% (inner line) of the spectrum log amplitude and show that man-made scenes contain more energy along the cardinal axis compared to the natural scenes. Images of (a) man-made (b) and natural scenes. Artificial images used for the adaptation based upon their relating power spectrum to emphasize (c) man-made (d) or natural image statistics. (e) Neutral adapter made up of circles and rectangles, combining man-made and natural power spectrum attributes.

predominance of contours oriented along the cardinal axes in man-made environments. They also point out that the statistics of orientation and scales are a good cue for scene categorization (Oliva & Torralba, 2001), and propose a simple linear model that uses the spectral principal components of these categories to allow semantic categorization between them (Torralba & Oliva, 2003).

While these studies document the presence and sufficient magnitude of statistical differences between images of natural and man-made environments, to date no psychophysical study has demonstrated that humans are able to exploit it for rapid scene categorization. Here we provide such a demonstration by documenting the presence of two aspects of human scene categorization that can be accounted for by a process that computes simple image statistics.

First, we test the effect of image inversions on performance because Fourier analysis is inversion-invariant due to the cardinal axes symmetry of the global frequency spectrum (Torralba & Oliva, 2003; see also Figure 1), i.e., upright and inverted images have identical image statistics and should therefore be equally distinguishable from other images.

Secondly, a scene categorization based on image statistics likely needs to be continuously calibrated, i.e., subjects probably categorize scenes into natural and man-made images by comparing a given scene's spectrum against an internal reference that represents an average of recent inputs. This would resemble similar processes in identity (Leopold, O'Toole, Vetter, & Blanz, 2001) or gender and race (Webster, Kaping, Mizokami, & Duhamel, 2004) categorizations based on images of faces. Such an approach is prone to the effects of adaptation, i.e., extended exposure to images stimulating those processing channels responsible for detecting extreme versions of one of the two categories should shift the subjects' categorization midpoint towards such adapters, if the adapted channels are indeed used in the categorization process.

In our experiments, subjects categorized greyscale environmental images in a two-alternative forced choice (man-made vs. natural) image rating task. We compared categorization performance for upright and inverted images of natural and man-made scenes and determined the effect of adapting with long-duration abstract stimuli that mimicked the prototypical orientation components of either man-made or natural scenes, respectively.

Our results show that performance was unaffected by image inversion and that the subjects' scene categorization was systematically affected by adaptation in line with the prediction sketched out above. Together the findings demonstrate that the human visual system exploits low-level image statistics for performing rapid scene categorization, an approach applicable for many categorization tasks and therefore probably widely employed.

METHODS

Twelve naive subjects (8 female and 4 male, ages 15–29) participated in the study. All subjects had normal or corrected-to-normal vision and gave written informed consent. Subjects sat in a dimly lit room, 57 cm from a computer monitor (85 Hz, 40 pixels/deg resolution) with their head stabilized on a chinrest. They were asked to categorize images briefly presented on a uniform grey background as man-made or natural scenes.

The test stimuli ("scene images") used were 316 grey level still images scaled to 13.3×10.9 deg (530×435 pixels) taken from the van Hateren and van der Schaaf Natural Stimuli Collection (1998). The images were selected from the collection such that about half of them were rated as man-made and half as natural by two of the authors with unlimited viewing time.

In each trial one test stimulus was presented for 12 ms between a spatial frequency adapting sequence and a mask stimulus (Figure 1). The mask (presented for 94 ms) appeared 94 ms after the test stimulus and was used to constrain the perceptual availability as a retinal afterimage. This inter-stimulus interval was chosen to be as short as possible and as long as necessary to allow acceptable performance.

The adapting stimuli were computer generated images of circles and/or rectangles that were composed such that they either matched the average power spectrum of all scene images (*neutral adapter*, made up of circles and rectangles), the spectrum of those scene images rated as man-made (*man-made adapter*, rectangles only), or that of the natural-scene images (*natural adapter*, circles only). A dynamic adaptation sequence of 10 adapting stimuli (117 ms each) was presented at the beginning of every trial. The adapting image sequence and the test images were separated by a 294 ms uniformly grey blank screen.

The three adapter types were used in separate experimental blocks of 316 test stimuli in a randomized sequence of 50% upright and 50% inverted images. In each block, each image was used upright with four subjects and inverted with another four subjects. Subjects were not told that inverted images were present. Each subject participated in two of the three adapting conditions, thus categorizing each image twice, once upright and once inverted. Results were analysed using standard Z-test for binomial distributions with adjusted p-values for multiple comparisons (Zar, 1999).

RESULTS

For each of the adaptation conditions each of the 316 test images was categorized four times in its upright and four times in its inverted orientation. For each image the number of "natural scene" responses was

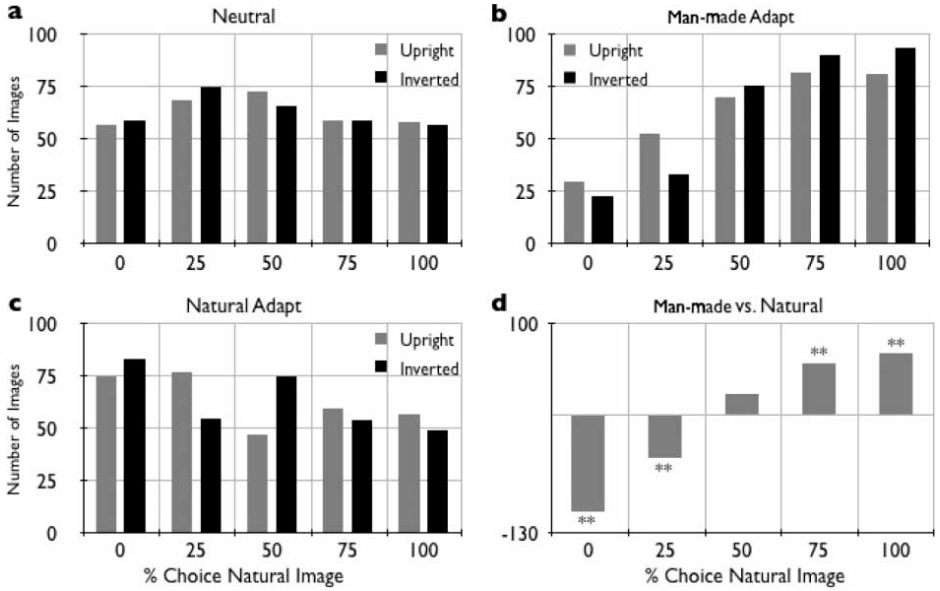


Figure 2. Histograms of number of images rated as man-made scenes (0%) by all four subjects that were shown a particular image, natural scenes (100%) or between, for (a) the neutral condition, (b) man-made-like adapters, and (c) natural statistics adapters. Categorization of upright and inverted images showed no significant difference throughout the three conditions (a–c), allowing to pool responses independent of orientation (d). Comparing man-made versus natural by subtracting the histograms show highly significant differences (d) (*$p_{\text{adjusted}} < .05$, **$p_{\text{adjusted}} < .01$).

counted across the four subjects that rated the image in the same orientation. For each possible count frequency (0, 25, 50, 75, and 100%) the number of images receiving the corresponding rating were counted (Figure 2a–c).

The light bars in Figure 2a show the resulting histogram for upright images in the neutral condition. The homogeneous distribution indicates that the subjects were able to perform the task, that the collection of images were not biased to one or the other category, and that the images varied as to their perceptual unambiguity. Comparing the response distribution against the one for the inverted images (dark bars) reveals no significant difference, indicating that the subjects could rate the inverted images just as well as the upright images.

Similarly, for the man-made and natural adapting conditions no significant differences were found for upright and inverted images. But the response distributions between these two adapting conditions were very different. Figure 2b shows that adaptation to the underlying statistics of man-made environments biased the categorization towards "natural" responses (see Figure 2b and 3b). A significant overall decrease ($Z = 4.21$, $p_{\text{adjusted}} < .01$ inverted, $Z = 3.08$, $p_{\text{adjusted}} < .05$ upright) of images collectively categorized as man-made (following adaptation to man-made image

statistics) produced a reshaped response to identify significantly ($Z = 3.37$, $p_{\text{adjusted}} < .01$ inverted) more natural aspects within the test images (Figure 2b). For the natural adaptation paradigm a strong opposite trend was present (see Figure 2c and 3b) and a direct comparison between the response distribution of man-made versus natural adapting stimuli revealed highly significant effects (Figure 2d, $Z > 7.74$, $p_{\text{adjusted}} < .01$, pooling over orientation).

DISCUSSION

Our data show that the human visual system is able to categorize novel environmental scenes rapidly and unaffected by inversion, indicating a neural mechanism not relaying on high-level image orientation dependent identification processes. This interpretation is supported by our finding that adaptation with an abstract image composed to mimic the orientation

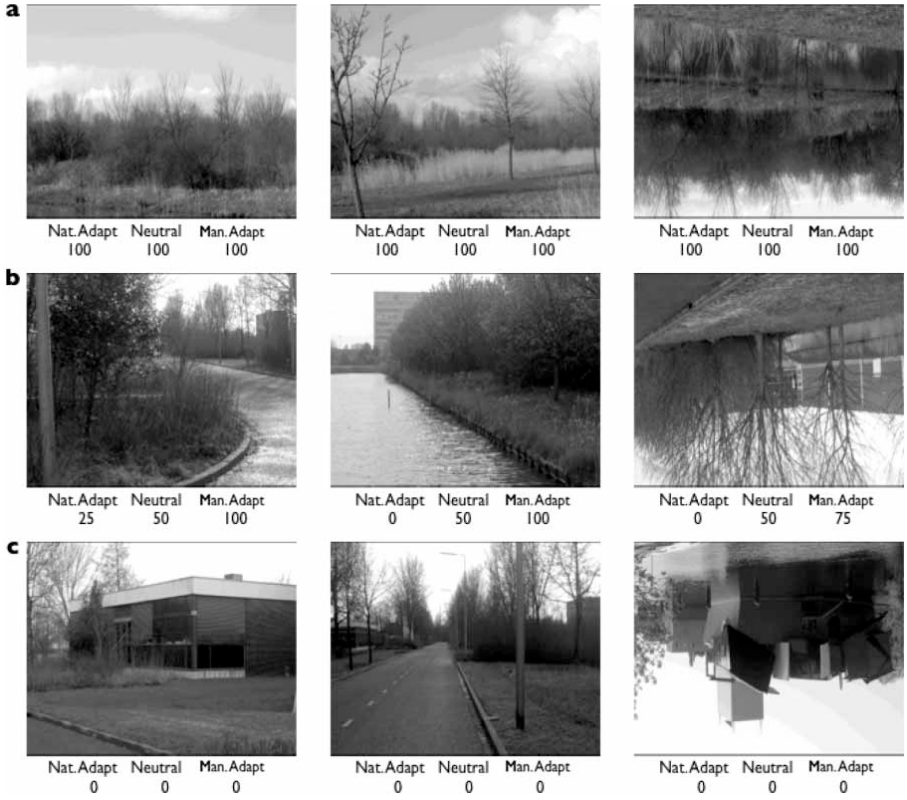


Figure 3. Illustration of natural scenes with their corresponding responses (below the image) in the three conditions. (a) Example of natural scenes unaffected by the adaptation to image statistics; (b) scenes judged to be ambiguous in the neutral condition shifted by the adapting conditions; (c) man-made scenes unaffected by the adaptation. Inverted images are shown in the last column.

content of a man-made scene biased subjects to report a given image as representing a natural scene more often than after exposure to an equally abstract adapting pattern mimicking the orientation composition of a natural scene (Figure 3d). This adaptation effect indicates that the abstract images affected specific processing channels that contribute to rapid scene categorization, documenting that the human visual system is not only highly sensitive to the statistical properties of the visual input but can also exploit patterns in those properties to perform such seemingly complex decisions as whether an image depicts a scene that is natural or man-made.

Two points need to be made when evaluating these findings: First, the rapid feedforward scene categorization process demonstrated by our findings is obviously just a first "best guess" of the visual system. It allows us to recover the "gist" of a scene (Braun, 2003). Scrutinizing the scene, if it remains visible (i.e., without masking), allows the visual system to employ its full range of object recognition systems resulting in a much more reliable categorization (Rosch, 1978) based on a fuller perceptual representation. Nevertheless, our data show a low-level scene analysis system that presumably operates on all inputs and might provide a preattentive screening for basic aspects of the visual signals entering cortex. As such the system could provide important input towards the construction of a saliency map of the visual environment (Treue, 2003).

Second, the approach employed by the visual system in extracting and interpreting the Fourier spectrum of the visual input is just one of many low-level analyses that can be performed by neuronal populations in the early visual system. Such systems could provide rapid estimates of many other categorical assessments of the visual input or even just patches of it.

In summary, our findings reveal a highly efficient system for constructing an internal representation of the visual input that relies on the feedforward extraction of "low-level" image features yet supports sophisticated perceptual judgements previously thought to require "high-level" image processing. This system appears to be particularly useful in case of high processing load, whenever fast judgements are needed and in animals that lack the sophisticated processing abilities of primate extrastriate cortex.

REFERENCES

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 213–253). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Braun, J. (2003). Natural scenes upset the visual applecart. *Trends in Cognitive Sciences*, *7*, 7–9.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, *16*, 152–160.

Kreiman, G., Koch, C., & Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neuroscience*, *3*, 946–953.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94.

Li, F. F., Van Rullen, R., Koch, C., & Rerona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, USA*, *99*, 9596–9601.

Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, *19*, 577–621.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*, 145–175.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509–522.

Rosch, E. (1978). Principles of categorization. In B. E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 28–49). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience*, *5*, 629–630.

Rousselet, G. A., Mace, M. J. M., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision*, *3*, 440–455.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139.

Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems*, *14*, 391–412.

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology*, *13*, 428–432.

Van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London Series B*, *265*, 359–366.

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception*, *30*, 655–668.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*, 557–561.

Zar, J. H. (1999). More on dichotomous variables. *Biostatistical analysis* (4th ed., pp. 555–558). Upper Saddle River, NJ: Prentice Hall.

Manuscript received January 2006
Manuscript accepted June 2006
First published online July 2006

II.II Adaptation to image statistics decreases sensitivity to the prevailing scene

The previous section (Adaptation to statistical properties of visual scenes biases rapid categorization 2.1) identified low-level statistical differences within environmental scenes to be sufficient information to characterize different types of environments. Categorization of rapidly displayed visual scenes could be strongly influenced by the statistical characteristics of the prevailing scenes. In the present study, we examined if the adaptation induced categorical shift between natural and man-made scenes describe distinct processing boundaries between these categories.

We hypothesize that the perceptual shift following the adaptation to statistical environmental properties during rapid image categorization results from suppressing information around the calibrated mean environment, thus only affecting the perception of scenes matching to the adapted environmental statistics.

To probe this prediction, we made use of a parallel processing paradigm displaying multiple images simultaneously. This method enabled us to analyze the exact modification of the categorization process, i.e. how each category (man-made or natural) is affected by adaptation to the statistics of a given environment. Adaptation to spatial frequency contents along different orientations adjusts the visual sensitivity according to the statistical spectrum of the adapted environment. This influenced the categorization of parallel processed scenes, corrupting accurate detection along with speedy processing of images only within the adapted category. These adjustments reveal a highly efficient processing mechanism within the visual system to rapidly extract category information as a result of removing redundant information to accentuate “low-level” statistical differences deviating from the mean. Further, our results reveal distinct processing boundaries between natural and man-made scenes, suggesting non-opponent processed categories.

Adapation to image statistics decreases sensitivity to the prevailing scene

Daniel Kaping, Stefan Treue
Cognitive Neuroscience Laboratory, German Primate Center, Goettingen, Germany
Bernstein Center for Computational Neuroscience, Goettingen, Germany

Abstract

Differences in the low-level image statistics of environmental scenes contain sufficient information to characterize different types of environments. Briefly displayed visual scenes can be strongly influenced by adaptation to the statistical characteristics of the prevailing visual input. We have previously reported an apparent processing boundary between basic categories of natural and man-made scenes. In the present study, we examine if this adaptation-induced bias in the categorization into natural vs. man-made scenes reflects distinct processing boundaries between these categories. During a rapid parallel multi-image detection task we singled out one target group's (man-made or natural) current state of categorization. Adaptation to spatial frequency contents along different orientations mimicking the overall statistical spectrum of a given environment adjusts human visual sensitivity only within the adapted category, influencing the categorization of parallel processed scenes, corrupting accurate detection along with processing speed. These category bound adjustments reveal a highly efficient processing mechanism within the visual system to rapidly extract category information as a result of removing redundant information to accentuate "low-level" statistical differences deviating from the mean. Our results reveal distinct processing boundaries between natural and man-made scenes, suggesting non-opponent processed categories.

Introduction

Rapid image categorization is the remarkable ability of the human visual system to extract sufficient information for categorical judgements of the visual scenes in a “wink” of time (Biederman, 1972; Potter, 1976). Correct classification of images with presentation times of 30 ms or less (Joubert, Rousselet, Fabre-Thorpe & Fize, 2009; Kaping, Tzvetanov & Treue, 2007; Fei-Fei, Iyer, Koch & Perona, 2007; Guyonneau, Kirchner & Thorpe, 2006; Kirchner & Thorpe, 2006; Thorpe, Fize & Marlot, 1996) along with category-specific brain activation within 150 ms of stimulus onset (Rousselet, Fabre-Thorpe & Thorpe 2002; VanRullen & Thorpe 2001) suggest the employment of a simple, feedforward processing mechanism. Such a system ought to base the rapid categorization of depicted scenes upon easily extractable global scene properties.

Torralba and Oliva (2003) have recently suggested a plausible method that would allow the visual system to carry out scene categorization without the need to first achieve full object recognition. Spatial frequency content along different orientations varies to an extent which permit man-made and natural images to be category specific processed, requiring only the presence of orientation-selective neurons such as those abundant in early visual cortex. Such a straightforward mechanism is supported by unchanged classification performance of inverted images (Kaping, Tzvetanov & Treue, 2007; Guyonneau, Kirchner & Thorpe, 2006; Drewes, Wichmann & Gegenfurtner, 2006; Rousselet, Mace & Fabre-Thorpe, 2003), and the low demand for attentional resources (Li et al., 2002) when the categorization of complex images is paired with attentionally demanding tasks. These findings imply a purely feedforward process (Koch & Tsuchiya, 2007) that is nevertheless able to deliver fast and reliable information about the scene at hand.

According to the general environmental input such a fast visual processing mechanism should be prone to rapidly adapt its response to optimize accurate detection of visually relevant changes. Using a two alternative forced choice task we previously showed a categorization bias following the adaptation to a dynamic sequence of computer generated images composed to match the power spectrum of natural or man-made environments (Kaping, Tzvetanov & Treue, 2007). Subsequently viewed ambiguous test images of environmental scenes were judged to be natural more often following the adaptation to man-made or judged to be man-made more often following the adaptation to natural image statistics. While these results provide the first evidence for a spatial frequency and orientation dependent scene classification mechanism, they did not

reveal man-made / natural category boundaries and the basis of the observed category shift. The sudden emergence of previously unattended features within a given environment may be based upon extracting deviating statistical content while disregarding the average statistical properties to an internal state calibrated according to the mean environment (Webster, Werner & Field, 2005). This adaptive adjustment optimizes information transfer of orientation selective filters by removing redundancy (Barlow, 1961) allowing the detection of features within the unadapted environment. Adaptation would therefore enable the visual system to shape a predictive code of the environment (Webster, 2005) creating a saliency map based upon differences from the overall surrounding environment (Treue, 2003). We hypothesize that natural and man-made scenes belong to distinct independently coded categories. The perceptual shift in rapid image categorization following the adaptation to statistical environmental properties results from suppressing information around the mean visual input, thus only affecting the perception of scenes matching the adapted environmental statistics.

To test this prediction, we made use of a modified parallel processing paradigm previously used by Rousselet, Thorpe and Fabre-Thorpe (2004). Multiple images were presented simultaneously (one target image among distractors) in an environment detection task following an adaptation mimicking the statistics of either man-made or natural environments. This method enabled us to analyze the exact modification of the categorization process; that is, how the categorization of each category (man-made or natural) is affected by adaptation to the statistics of a given environment.

In two separate complementary experiments, one presenting two parallel streams of stimuli and the other presenting four parallel streams (to exclude location biases), we determined the change of correct environment detection and the associated change in reaction time (RT). Both experimental conditions were subdivided in a no-adapt (baseline) condition, adaptation to statistical properties of natural (rural) environments and the adaptation to the statistical properties underlying man-made (urban) environments. To exclude location biases; subjects performed a two / four - alternative forced choice, category detection task of either man-made or natural environments.

Methods

Twenty-four naive subjects (ages 19 – 31, 16 men) participated in the study. All subjects had normal or corrected-to-normal vision and gave written informed consent. Subjects sat in a dimly lit room, 57 cm from a computer monitor (85 Hz, 40 pixels/deg resolution) with their head stabilized against a headband and resting on a chin plate. They were

instructed to identify, as fast as possible, the position of a test image among distractor images during a man-made or natural scene detection task. The test and distractor images (environmental scenes) used, were 600 grey level still images of 10 * 10 deg (400 by 400 pixels). The images were selected from the collection such that they corresponded with the designated category (i.e., man-made or natural) and such that their respective power spectrums matched their category according to Torralba & Oliva, 2003.

Two-Image Condition

In each trial, one test image of a given environment (i.e., man-made or natural) was presented simultaneously with one distractor scene of the opposite category for 12 ms followed by a dynamic visual mask. The inter-stimulus interval between test and mask sequence was set to 96 ms, chosen to be as short as possible and as long as necessary to allow acceptable performance. The mask (presented for 94 ms) was used to constrain the perceptual availability of the test sequence (see Fig. 1, illustration of the time course of the experiment).

The experiment included a total of three blocked conditions (eight subjects per condition): one adapting to man-made image statistics, one to natural spectra, and one non-adapting condition (baseline). During the adaptation conditions stimuli preceding the test series were computer generated images of circles or rectangles that were composed such that they either matched the power spectrum of scene images rated as man-made (*man-made adapter*, rectangles only) or those of the natural-scene images (*natural adapter*, circles only) (see Kaping et al., 2007). A dynamic image sequence of ten adapting stimuli (117 ms each) was presented at the beginning of each adaptation trial. The adapting image sequence and the test sequence depicting the environmental images was separated by 294 ms of uniformly grey screen to ensure unhindered view and onset recognition of the briefly presented test sequence (see Fig. 1a).

Each adapter type (man-made or natural) was used twice, once when detecting the man-made environmental image and once with the natural environmental image. Eight subjects were tested per adaptation condition and were randomly assigned into separate environment detection task settings (e.g. man-made adapt to detect man-made, man-made adapt to detect natural; natural adapt and no adapt followed the same scheme). During each block of 100 trials, 50 target images were presented on the left and 50 target images on the right (one target and one distractor simultaneously, position chosen randomly for each trial). Subjects, fixating on a central fixation spot, were

instructed to answer as fast as possible on which side of the fixation spot (left or right) the target environment was displayed. In the third, no-adapt condition, an additional eight subjects carried out the environment detection task in the absence of the adaptation sequence. Test and distractor images appeared 294 ms after trial onset followed by the previously described mask cycle. To prevent learning and the recognition of individual images based upon object properties, each test image of a given target environment was only presented as a test image once; but could reappear as a distractor image when the other environment was targeted. We forwent presenting two images intra-hemifield as Rousselet et al. (2004) observed no difference between inter- and intra-hemifield presentation of two simultaneously displayed images. To control the possibility that subjects did not apply a single image detection strategy inferring the correct category location through exploiting only one presentation side (left or right) or one of the two categories, a four image detection task was introduced. This ensures the utilization of parallel processing as a result of “forced” handling the increased number of stimuli in absence of increasing processing time.

Four-Image Condition

The four-image condition followed the same experimental procedure as the previously described two-image condition (divided into three settings: no-adapt, natural adapt, man-made adapt, presenting a total of 400 images per condition). The display was divided into quadrants with each containing one image (one target and three distractor images) with positions chosen randomly (see Fig. 1b). Subjects were directed to answer as fast as possible in a four alternative forced choice reaction time task in which of the four (upper-left, upper-right, lower-left, lower-right) quadrants the target environment was shown while fixating the center of the screen. Each test sequence consists of four simultaneous and rapidly (12 ms) displayed environmental photographs followed by the dynamic mask sequence. Depending upon the overall condition, either an adaptation sequence of images mimicking man-made or natural environmental statistics preceded the test sequence, or either a gray blank screen (no adapt condition) was displayed.

Similarly to previous studies we compared performance of different adaptors between different subgroups of subjects. That is, subjects alternated adapters in-between the two- and four image condition (e.g. two-image condition adapt man-made, four-image condition adapt natural and vice versa).

To assess the influence of a given adapting sequence mimicking statistical environmental content and the resulting visual adjustment RT and correct localization of the target environment were analyzed. The state of adaptation could both influence the correct recognition of an environmental scene and impact RT's of an observer rapidly categorizing environmental images. Increased RT's correlated with wrongful environmental image detection could be related to higher processing requirements (Pins and Bonnet, 1996) resulting from the adaptive adjustments made by the visual system evoked by the preceding environmental statistics. The observed relationship between correct trial outcome and RT's (Thorpe, Fize & Marlot, 1996; Pins and Bonnet, 1996) promotes RT as a sensitive analysis tool in this adaptation-influenced rapid categorization task. We therefore analyzed the performance of natural and man-made environmental image detection influenced by different adaptational states of the observer and the associated RT's.

The percent correct were analyzed with a one-way analysis of variance adjusting for multiple comparisons (Tukey's Test) and RTs distributions with Kruskal-Wallis non-parametric test (multiple comparisons with Dunn's test).

Results

Experiment 1: Category detection within two parallel dynamic streams

In the two-image condition, with no adapting sequence subjects showed no significant difference between category types in locating the target image in the presence of one distractor image. The task required the subject to respond as fast and accurate as possible indicating the correct location of the man-made environment while a natural distractor image was present, or to respond to the natural image disregarding the man-made distractor. Subjects had no difficulty identifying the correct category position (Fig. 2; man-made 96.5 % correct, natural 95.25 % correct). These findings are consistent with results reported by Fei-Fei et al. (2007) who obtained no differences perceiving man-made outdoor over natural outdoor images in a rapid image content recognition paradigm with varying presentation times. Our categorization results in the no-adapt condition stand in contrast to Rousselet et al. (2004), where subjects reported only 75% correct in a parallel two-image animal detection task. This results from task differences in that our subjects categorized images at an earlier more "basic-level" (Rosch, 1978). RT's values did not differ for man-made (median 465 ms, mean 440 ms, minimum 260 ms for correct image detection) and natural (median 482 ms, mean 460 ms, minimum

280 ms for correct image detection) two-image no-adapt condition (Fig. 2; Kruskal-Wallis non-parametric test of RT distributions, $p>0.05$).

When subjects were instructed to detect the Man-made environment, the two adaptation conditions showed different effects in categorization accuracy. Adapting to man-made environmental statistics decreased subjects' performance compared to the no-adapt condition (82.25% vs 96.5%), whereas adapting to natural environments did not modify their performance 97% vs 96.5%)(Fig. 2). This result was confirmed by a one-way ANOVA (three levels: no-adapt, Man- and Natural-adapt) demonstrating a main effect of the adapting condition ($p<0.01$). Post hoc comparisons of the means showed that percent correct for the man-made adaptation condition, detecting the man-made environment was significantly lower than in the two other conditions ($p<0.01$). The percent correct changes were oppositely related with changes in subjects' RTs. RTs for detecting man-made environment were higher when subjects were adapted to the statistics of man-made environments compared to the no-adapt condition (medians: 674 vs 440 ms), and slightly higher for natural environment adaptation compared to no-adaptation conditions (490 vs 440 ms) (Fig.2). This was confirmed by a Kruskal-Wallis non-parametric rank test on the RT distributions that showed a main effect ($p<0.01$). Post hoc multiple comparisons demonstrated that all three RT distributions were different ($p<0.01$).

The above results for detecting the man-made environment within two-parallel dynamic streams of images were similar for the natural environment detection task. Adaptation to the statistics of natural environments seemed to decrease the percent correct detection when compared to the no-adaptation condition (92.25% vs 95.25%), while adaptation to man-made environmental statistics present a small increase (97% vs 95.25%) in correct natural image detection. These observations were confirmed by the one-way ANOVA that demonstrated a main effect of adapting condition ($p<0.05$). Multiple comparisons showed that correct answers in condition with adaptation to man-made statistics was significantly different from the condition with adaptation to natural environment statistics ($p<0.05$), whereas the two other comparisons did not reach significance level ($p>0.05$). Adaptation to natural statistics influenced strongly the RTs. Subjects' RTs where much higher when compared to the remaining two conditions (medians; natural-adapt 500 ms, man-made adapt 450 ms, no-adapt, 460 ms). This observation was confirmed by the Kruskal-Wallis statistical test on the RT distributions (main effect: $p<0.01$). Post hoc

multiple comparisons confirmed that the natural-adapt RT distribution was significantly different from the remaining two ($p<0.01$), whereas RTs for man-made adaptation were not significantly different from the no-adapt condition.

Experiment 2: Category detection within four parallel dynamic streams

As well as in the no-adapt two-image conditions, the no-adapt four-image conditions (detect man-made or natural) revealed no significant differences between category detection (Fig. 3; man-made detect 74.75%, natural detect 75.25%). Resulting form the increased number of stimuli the overall performance of the four-image conditions was significantly lower than in the two-image conditions (~20%). The increasing difficulty of accurate image detection in conjunction with increasing total number of simultaneously presented images has previously been reported by Rousselet et al. (2004).

For detecting man-made environments in the four parallel streams, introduction of the man-made adaptation sequence decreased correct image localization (Fig. 3), whereas natural-adaptation sequence did not modify subjects' performance. This was confirmed by the one-way ANOVA (3 levels) with a highly significant effect ($p<0.01$). Post hoc multiple comparisons confirmed a statistically significant decrease of the percent correct in the man-made adaptation condition when compared to the remaining two conditions ($p<0.01$), and no significant difference between natural-adapt and no-adapt conditions. The man-made adaptation condition also increased overall processing time of images belonging to the same category (Fig. 3) compared to the two other conditions (median RTs: man-made adapt 780 ms, natural-adapt 620 ms, no-adapt 635 ms). This effect was confirmed by the Kruskal-Wallis test on RTs distributions ($p<0.01$), and post-hoc multiple comparisons revealed that all three RT distributions were significantly different (man- vs no-adapt, $p<0.01$; man- vs natural-adapt, $p<0.01$; no- vs natural-adapt, $p<0.05$).

Detection of natural environments in the four-image parallel streams was also affected by introducing the two adaptation sequences (Fig. 3). Natural environment adaptation decreased detection performance in comparison to no-adaptation (66.75% vs 75.25%), whereas man-made adaptation sequences almost did not modify subjects' performance (77.25% vs 75.25%). The one-way ANOVA reported an effect of adaptation condition ($p<0.01$). Post-hoc multiple comparisons revealed a statistical difference only between man- and natural-adapt conditions ($p<0.05$). In contrast, RTs were strongly affected by

the different conditions demonstrating an increase in RTs for the natural adaptation condition (medians; natural-adapt 730 ms, man-made adapt 605 ms, no-adapt 687 ms). The RTs distributions were shown to be statistically different (Kruskal-Wallis test, $p<0.001$) and post-hoc multiple comparisons demonstrated statistical differences for natural- vs no-adapt and natural- vs man-made-adapt (both, $p<0.01$).

No significant position dependent effect can be reported. Observer's performance was unaffected not only for left right test image position but also for vertical / horizontal positioning.

Discussion & Conclusion

The present results show two critical aspects of visually adaptive adjustments along with rapid scene detection / categorization: one providing further evidence of a suitable "low level" environmental calibration mechanism and another of removing redundant information based upon adaptive adjustments to the overall environmental statistics. Importantly, through comparing the performance of parallel image detection under three possible conditions (no-adapt, man-made adapt & natural adapt) we provide further insight to the spatial frequency / orientation dependent fast image recognition process. Subjects appear to have average environmental categories internally represented not only based upon possible content but also based upon purely underlying statistical properties. The disruption of overall performance following the adaptation to the task relevant environmental statistics can be understood as the aftereffect commonly used to describe different post-adaptation perception. Here it appears that man-made and natural environmental statistics produce different aftereffects and therefore are processed differently.

While our man-made adapters produce a reduction in correct man-made image detection and an increase in RTs for both two-image and four-image presentations, the natural adapters produced significant changes only in the RTs but detection performance for natural images was not significantly impaired. We interpret this discrepancy by the chosen circular shape in the Fourier domain of our natural adapter, which does not simulate all possible natural environments. The exact Fourier template of natural environments might be a more complex combination of orientations and spatial frequencies than the one we chose, because of the wider range of Fourier shapes for natural environments compared to man-made environments (see Figure 3 in

Torralba and Oliva, 2003). Therefore, it appears to us that the RT measure reveals the adaptor effect and contributes as an important parallel measure to detection/ categorization performance (e.g. Pins, Bonnet, and Dresp, 1999). Thus, despite the natural adapter's smaller efficiency unfortunately not targeting the whole natural images' Fourier spectrum its effect can still be observed.

With our paradigm we were able to modify observers performance on a given image detection/categorization task through adaptation to the task relevant environmental statistics. The low-level image features give rise to a stimulus-related pre-recognition (Johnson & Olshausen, 2003) before complex stimulus features trigger object based scene recognition. This allows the visual system to channel a given input based upon fast extractable features. Further higher-level processing could be computed in a feed-forward manner, and thus used for forming simple templates based on combinations of spatial frequencies and orientations for detection, categorization or classification of images at superordinate levels (vanRullen and Thorpe, 2001; Torralba and Oliva, 2001, 2003; Viéville and Crahay, 2004). Joubert et al., 2009 support the hypothesis that the amplitude spectrum serves the visual system as an early cue and provide evidence the phase spectrum emphasized by Loschky and Larson (2008) , Wichmann, Braun & Gegenfurtner (2006) is used to enrich an early categorization. Our adaptation-based results strongly support the idea that human subjects are able to categorize rapidly with a single wave of feed-forward activity (see also Joubert et al., 2009, VanRullen and Thorpe, 2001; VanRullen, Delorme, Thorpe, 2001).

We would like to emphasize that our results are clearly showing that this decision process is based on templates of different spatial-frequencies at different orientations, and not only due to some modification of early visual processing. This is seen in our "natural" adaptor that was created for symmetrically adapting spatial frequencies to all orientations (see Fig. 4, circular shape of the Fourier power spectrum). Therefore, if the only effect of this adaptor was to modify the early transmission of spatial frequencies at different orientations, it would have changed both categorization task in exactly the same manner. We clearly observed a differential effect between man-made and natural categorization in performance and RTs that can only be explained with an adaptation to the full Fourier power spectrum template. This result confirms and extends our initial finding (Kaping, Tzvetanov, Treue, 2007) by demonstrating that Fourier power spectrum templates are a basis for performing categorization task of environmental scenes and

that they are to some extent independently processed.

While two mechanisms are plausible for aiding the observer to achieve correct categorization: one allowing the system to code a given environmental change based upon increased sensitivity away form the calibrated mean environment therefore attracting and requiring more attentional resources – the other based upon decreasing sensitivity to the mean demanding less change in coding properties to a presented change; the later seems to apply. Previously unattended image features are accented through repressing redundant information following the adaptation process assigning the neutral point of visual coding. This adaptational shift could allow a modulation of the visual saliency map (Treue, 2003) suggesting a close relationship between adaptation and attention (Boynton, 2004). A recent study (David, Hayden, Mazer, Gallant, 2008) in macaque area V4, critical for form and shape perception, suggests attention mediated filter properties of individual neurons for spatial frequency and orientation in a natural scenes match to sample task.

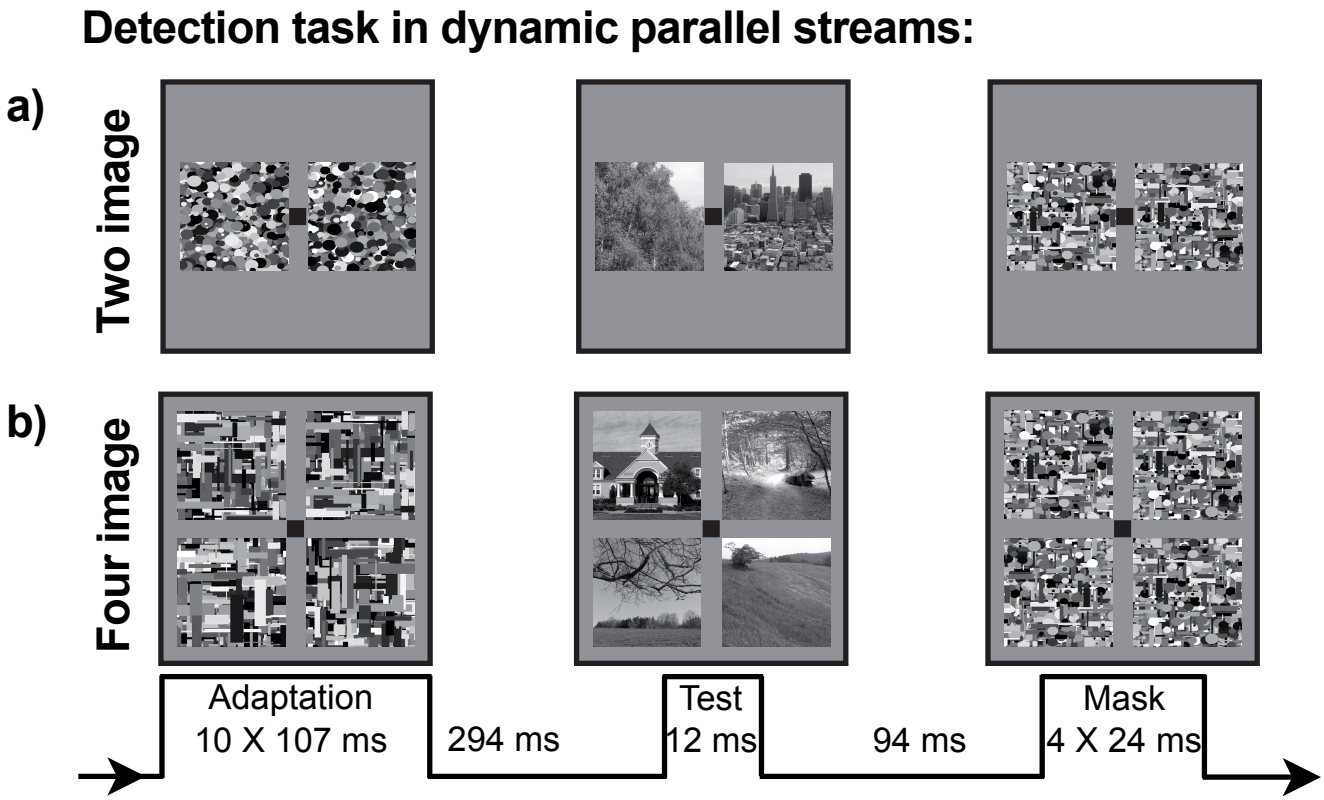


Figure 1. Schematics of the events in a two image detection trial (a): 10 dynamically changing adaptor images (here mimicking the power spectrum of natural scenes see Fig. 4) presented for 107 ms each followed by a blank period. Two simultaneously presented test images of opposing categories one serving as target the other as distractor followed by a brief blank period and four masking images made up of a mix of natural and man-made adapters; four image detection trial (b): here presented with a man-made adaptor sequence following the same time course as the two image trial (a) presenting four test images simultaneously for 12 ms one target and three distractors.

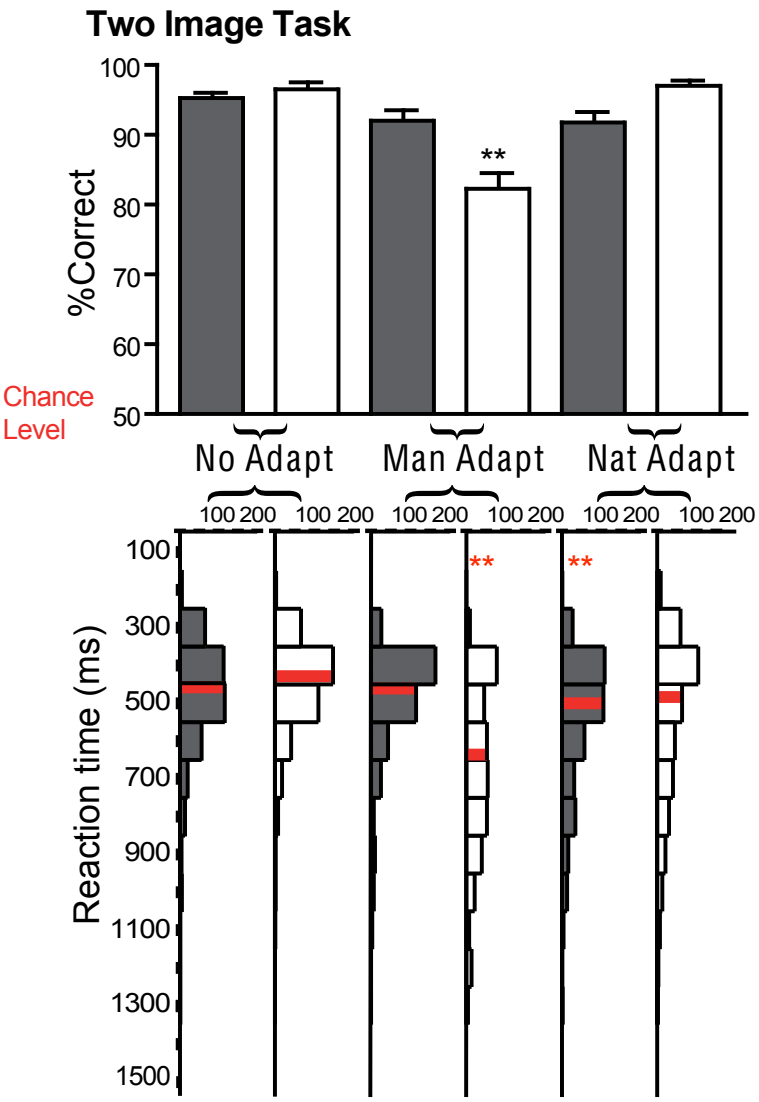


Figure 2. Top bar-graph - Average accuracy percent correct of natural (gray) / man-made (white) responses in the presence of a single opposing category distractor image as a function of preceding adaptor stimulus. Error bars indicate the standard error of the mean across subjects. Bottom bar-graph - Average median RT for correct responses (ms).

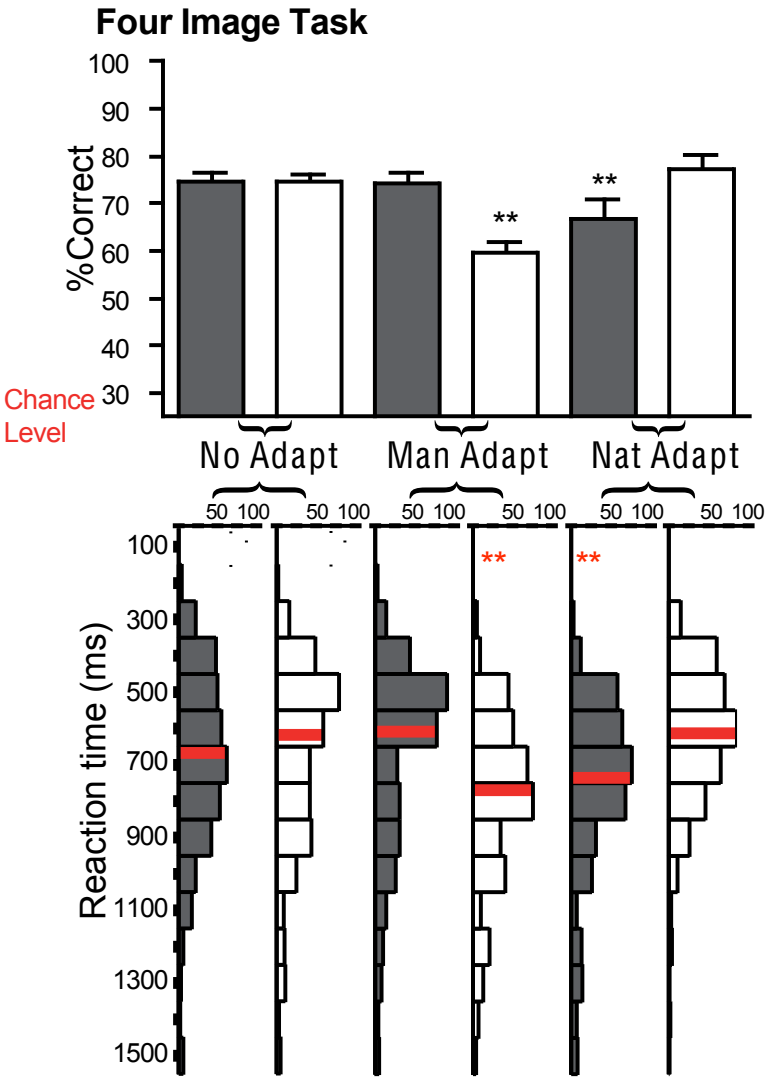


Figure 3. Top bar-graph - Average accuracy percent correct of natural (gray) / man-made (white) responses in the presence of a three opposing category distractor images as a function of preceding adaptor stimulus. Error bars indicate the standard error of the mean across subjects. Bottom bar-graph - Average median RT for correct responses (ms).

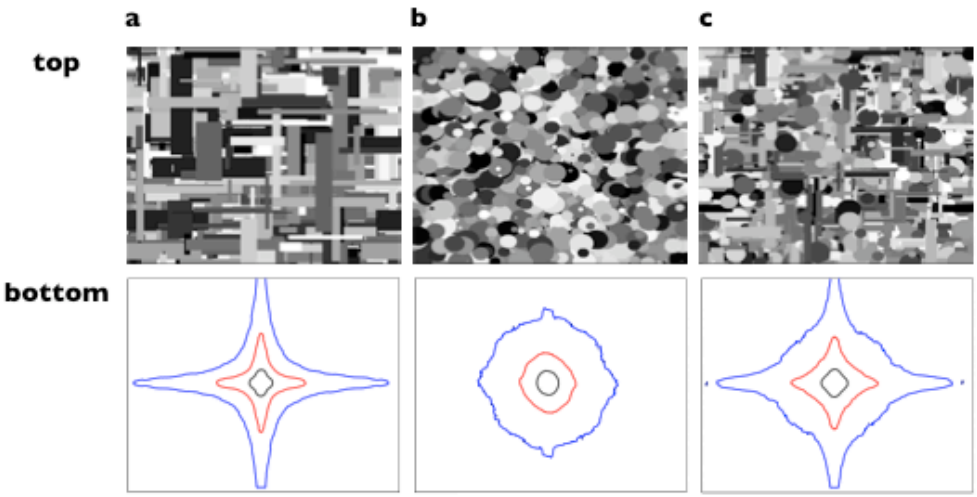


Figure 4. Artificial images used for used for the adaption based upon their relating power spectrum (a) man-made, (b) natural and an example of a masking image (c) composed of circles and rectangles, combining man-made and natural power spectrum attributes.

References

Barlow HB (1961). Possible principles underlying the transformation of sensory messages. In Sensory Communication, ed. WA Rosenblith, pp. 217–34. Cambridge, MA: MIT Press

Biederman I (1972). Perceiving real-world scenes. Science; 177(43):77-80

Boynton GM (2004). Adaptation and attentional selection. Nat Neurosci.;7(1):8-10

Clifford C W G, Webster M A, Stanley G B, Stocker A A, Kohn A, Sharpee T O and Schwartz O (2007). Visual adaptation: Neural, psychological and computational aspects. Vision Research, 47: 3125–3131

David SV, Hayden BY, Mazer JA, Gallant JL (2008). Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. Neuron.;59(3):509-21.

Drewes, J, Wichmann FA and Gegenfurtner K (2006). Classification of natural scenes: Critical features revisited. J Vis.;6(6):561

Fei-Fei L, Iyer A, Koch C, Perona P (2007). What do we perceive in a glance of a real-world scene? J Vis.;7(1):10

Guyonneau R, Kirchner H, Thorpe SJ (2006). Animals roll around the clock: the rotation invariance of ultrarapid visual processing. J Vis.;6(10):1008-17

Johnson JS, Olshausen BA (2003). Timecourse of neural signatures of object recognition. J Vis. 2003;3(7):499-512

Joubert OR, Rousselet GA, Fabre-Thorpe M, Fize D (2009). Rapid visual categorization of natural scene contexts with equalized amplitude spectrum and increasing phase noise. J Vis.; 9(1):2.1-16

Kaping D, Tzvetanov T, Treue T (2007). Adaptation to statistical properties of visual scenes biases rapid categorization. Visual Cognition; 15: 12-19

Kirchner H, Thorpe SJ (2006). Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vision Res.;46(11):1762-76

Koch C, Tsuchiya N (2007). Attention and consciousness: two distinct brain processes. Trends Cogn Sci.;11(1):16-22

Li FF, VanRullen R, Koch C and Perona P (2002) Rapid natural scene categorization in the near absence of attention. Proc. Natl. Acad. Sci.; 99, 8378 - 8383

Loschky LC and Larson AM (2008). Localized information is necessary for scene categorization, including the Natural/Man-made distinction. J Vis.;8(1):4.1-9

Oliva A and Torralba A (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. International Journal of Computer Vision;42(3), 145–175

Pins D, Bonnet C (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. Percept Psychophys.r;58(3):390-400

Pins D, Bonnet C, Dresp B (1999).Response times to Ehrenstein illusions of varying subjective magnitude: complementarity of psychophysical measures. Psychon Bull Rev; 6(3):437-44

Potter MC (1976). Short-term conceptual memory for pictures.
J Exp Psychol Hum Learn; 2(5): 509-22

Rosch, E (1978). Principles of categorization. In E. Rosch, & B.B. Lloyd (Eds.), Cognition and categorization. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rousselet G, Fabre-Thorpe M & Thorpe SJ (2002). Parallel processing in high level categorization of natural images Nature Neuroscience, 5, 629-630

Rousselet GA, Macé MJ, Fabre-Thorpe M (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. J Vis.;3(6):440-55

Rousselet GA, Thorpe SJ, Fabre-Thorpe M (2004). Processing of one, two or four natural scenes in humans: the limits of parallelism. Vision Res.;44(9):877-94

Thorpe S, Fize D, & Marlot C (1996). Speed of processing in the human visual system. Nature, 381, 520-522

Torralba A, Oliva A (2003). Statistics of natural image categories. Network.; 14(3): 391-412.

Treue S (2003). Visual attention: the where, what, how and why of saliency. Curr Opin Neurobiol.;13(4):428-32

VanRullen R, Delorme A and Thorpe SJ (2001). Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. Neurocomputing, 38-40(1-4), 1003-1009

VanRullen R, Thorpe SJ (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. Perception; 30(6):655-68

Viéville T. and Crahay S (2004). A deterministic biologically plausible classifier. Neurocomputing.; 58-60, 923-928

Webster MA, Werner JS, and Field DJ (2005). Adaptation and the phenomenology of perception. In C. W. G. Clifford & G. Rhodes (Eds.), Fitting the mind to the world: Adaptation and aftereffects in high level vision (pp. 241–277). Oxford: Oxford University Press.

Wichmann FA, Braun DI and Gegenfurtner KR (2006). Phase noise and the classification of natural images. Vision Res.;46(8-8):1520-1529

II.III The face distortion aftereffect reveal norm-based coding in human face perception

Faces represent the most relevant, everyday present stimulus within our visual word. Despite the physical similarities of faces, we are able to distinguish subtle differences along a large array of possibilities. The difficulty to discriminate between faces becomes more evident when we look at unfamiliar faces from a different race. This has been described as the “other-race-effect” and is not restricted to the identification of other-race faces. Rather, it describes the difficulty we possess in discriminating faces outside of familiar categories. One approach to explain our ability to discriminate faces is a multidimensional face space. Centered around a neutral average face, faces are placed along vectors in this multidimensional space. The axes represent the variation in features away from the average prototype face, while the distance between faces corresponds to their similarity. The prototype is not set within a rigid space but it can be shifted according to the exposure to different faces. Face perception is known to be strongly affected by adaptation to previously viewed faces. These adaptation effects may play an important role in calibrating face representations in the brain. However, this recalibration remains poorly understood, and could reflect a recentering of face coding around a set norm. We examined the neural correlates of face adaptation by monitoring haemodynamic activity with fMRI while observers viewed alternations between a normal and distorted face. Contrasting the switch of normal and contracted faces elicited strong activity in both the face sensitive fusiform gyrus and superior temporal sulcus, a region involved in the perception of changeable aspects of faces. More surprisingly, it also modulated responses in the motion sensitive area hMT+, consistent with a “motion aftereffect” to the illusory motion in the static images as their perceived shape changed / renormalized over time. However, these response changes were asymmetric with increased responses to normal faces following the adaptation to the distorted face than vice versa. This asymmetry parallels the relative salience of the perceptual aftereffects suggesting adaptation to influence the representations of faces relative to a set norm, providing evidence of norm-based face coding.

fMRI asymmetries and a novel motion area activation support a norm-based code in human face perception

Daniel Kaping[1,3], Carmen Morawetz[1,2,], Juergen Baudewig[2], Stefan Treue[1,3], Mike Webster[4], Peter Dechent[2]

[1]Cognitive Neuroscience Laboratory, German Primate Center, Goettingen, Germany

[2]MR-Research in Neurology & Psychiatry, Medical Faculty, Georg-August University Goettingen, Germany

[3]Bernstein Center for Computational Neuroscience, Goettingen, Germany

[4]Department of Psychology, University of Nevada, Reno, NV, USA

Summary

Humans are able to distinguish thousands of faces. Encoding individual faces by their deviation from a norm or average face would be an effective code underlying this perceptual ability. Norm-based coding is supported by behavioral studies that have examined how face perception is affected by prior adaptation to faces. We probed the neural correlates of these perceptual aftereffects for faces using functional magnetic resonance imaging (fMRI) to measure responses to normal faces after adapting to abnormal (distorted) faces, and vice versa. Haemodynamic responses were much stronger for the aftereffects induced in a normal face by a distorted adapting face. This asymmetry suggests that normal faces are represented by neutral response states, consistent with a norm-based code in face-selective cortical areas. Adaptation causes such powerful changes in the encoding of test faces that we observe a BOLD response in the motion-sensitive medial temporal cortex (hMT+) most likely caused by the illusory deformation of the test face as the aftereffect dissipates with time. This fMRI signal is a novel form of activation of motion sensitive areas as it is not caused by previous motion nor by stimuli that are associated with movement.

Keywords: face, adaption, fMRI, perceptual aftereffect, visual cortex

Introduction

Face perception can be strongly affected by adaptation to previously viewed faces. For example, after adapting to a contracted face, a normal face appears expanded (Webster and MacLin 1999; Watson and Clifford 2003). Adaptation aftereffects have been demonstrated along many of the dimensions that characterize natural variations in faces, including identity, expression, gender, or ethnicity (Leopold, O'Toole et al. 2001; Rhodes, Jeffery et al. 2003; Webster, Kaping et al. 2004; Rhodes and Jeffery 2006). These face aftereffects follow timecourses of logarithmic build-up and exponential decay (Leopold, Rhodes et al. 2005; Rhodes, Jeffery et al. 2007) that resemble more “low-level” adaptational adjustments and thus probably depend on similar processes of adaptation. However, unlike low-level aftereffects, face aftereffects transfer across size (Leopold, O'Toole et al. 2001; Zhao and Chubb 2001), position, and head orientation (Watson and Clifford 2003), implicating sensitivity changes in high-level neural mechanisms.

Adaptation-induced changes in face perception have been utilized to infer underlying coding mechanisms (Robbins, McKone & Edwards, 2007, Rhodes, Jeffery, Watson, Clifford, Nakayama, 2003) and provide an important test between alternative models of face coding. In norm-based codes, the dimensions characterizing a face are represented relative to a neutral prototype (Fig 1a). Because this prototype corresponds to a null response, adaptation to it does not bias sensitivity and thus does not alter the appearance of other faces. Conversely, adapting to a non-average face reduces the response to the adapting face (Fig 1b). This shifts the norm or neutral point toward the adapting face, inducing a negative perceptual aftereffect. An alternative to norm-based codes is an exemplar code in which faces are represented within multiple mechanisms tuned to narrow levels along the dimension and thus to particular features or configurations (Fig 1c). In this model no facial configuration is special, and adapting to any face should reduce the response to that face and bias the appearance of similar faces (Fig 1d). The two models thus make different predictions for asymmetries in the adaptation-induced aftereffects. Previous behavioral studies have found that adaptation alters appearance relative to a norm (Leopold, O'Toole et al. 2001; Robbins, McKone and Edwards, 2007) and shows a strong asymmetry in the aftereffects for normal or distorted adapting faces (Webster and McLin, 1999), pointing to a norm-based code. However, the neural correlates of these response changes have not been explored.

Here we tested the neural substrate of the perceptual aftereffects of face adaptation by monitoring the haemodynamic response in face-selective cortical areas while observers viewed an alternation between an original and distorted face using fMRI. This experimental design allowed us to readily compare the modulation of activity that accompanies the perceptual aftereffects for pairs of faces that varied in the degree to which they differ from a potential prototype. The perceptual aftereffects include an illusory deformation movement in the test face as the aftereffect dissipates with time. We asked whether this powerful correlate of the rules of face encoding would trigger neural changes outside face coding areas, by measuring the haemodynamic response in cortical areas coding motion.

Materials and Methods

Subjects

Thirteen right-handed, healthy volunteers (4 male; mean age ± SD: 27 ± 5 years) with normal or corrected to normal vision participated in the study. All subjects gave written informed consent to participate in the study, which was approved by the local ethics committee.

Localizer experiments

Stimuli

Face stimuli were grey-scale full-front digital images of six young males and six young females (Kovács, Zimmer et al. 2006). The photographs had no obvious gender-specific features, such as facial hair, jewelry, glasses or make-up. They were fit behind an oval mask (fit into a square of 400 x 400 pixels, 7.3° of height) eliminating the outer contours of the faces. House stimuli were grey-scale full-front images of fifteen different houses, which were fit behind the same oval mask as the face stimuli. Stimuli were presented in the centre of the screen on LCD googles (Resonance Technology, Northridge, USA) using the stimulation software Presentation (Version 9.00, Neurobehavioral Systems, USA).

Procedure

The fusiform face area (FFA; left: x = -40, y = -50, z = -14; size = 3102 voxel; right: x = 42, y = -52, z = -14; size = 9109 voxel) in the fusiform gyrus and the superior temporal sulcus (STS; left: x = -48, y = -51, z = 12; size = 3021 voxel; right: x = 50, y = -45, z = 9; size = 3680 voxel) were localized by presenting 18 sec blocks of grayscale face or house

images interleaved with 18 sec of images with the same amplitude spectra but scrambled phase spectra. Each face and house block was repeated six times. The coordinates of the FFA and STS are consistent with the peaks of activation reported previously (Kanwisher, McDermott et al. 1997; Ishai et al., 2005).

To functionally determine the motion-specific human medial temporal area a movie sequence mimicking the illusory motion apparent during the face distortion aftereffect was used lasting for two seconds. Following the adaptation to a contracted face subjects viewed a slightly expanded version of the original face that changed in a contracting motion to the original face. Stimuli were presented in a blocked design (17x18 sec cycles) starting with a contracted face and then altering between the original and contracted face of the same identity. Two localizer runs were performed using faces of different identities. Subjects were instructed to view the face stimuli passively by fixating the middle of the screen. On the basis of the obtained activation map contrasting the movie sequences against each other and the mean coordinates of the hMT+ (left: x = -46, y = -66, z = 3; right: x = 43, y = -63, z = 3) obtained from the Brede Database (Nielsen, 2003), a mixed functional and anatomical ROI in form of a cuboid (14x14x9mmm³; number of voxels: left: 1917; right: 2221) was defined.

fMRI Data Acquisition

MRI was performed at 3 Tesla (Magnetom Trio, Siemens, Erlangen, Germany). Initially, a high-resolution 3D T1-weighted anatomical dataset was acquired for each subject (176 sagittal sections, 1x1x1mm³). For fMRI a T2*-weighted, gradient-echo echo planar imaging technique recording 22 sections of 4mm thickness oriented roughly parallel to the calcarine sulcus at an in-plane resolution of 2x2mm² was used (repetition time = 2000ms; echo time = 36ms; field-of-view = 192x256mm²). This resulted in 153 whole brain volumes in each time series. Three time series were obtained in each subject in a single fMRI session (one series to determine the FFA and STS and two series to identify hMT+). The order of the series was randomized.

fMRI Data Analysis

Data analysis was performed with BrainVoyager QX (Brain Innovation, Maastricht, The Netherlands). Preprocessing included 3D-motion-correction, temporal high pass filtering (3 cycles/run), linear trend removal, spatial smoothing (Gaussian smoothing kernel, 4 mm full width half maximum) and transformation into the space of Talairach and Tournoux (Talairach & Tournoux 1988). Regions of interest (ROIs), namely the FFA, STS and hMT

+, were defined on the basis of group analysis using a false discovery rate of q(FDR)<0.01. The contrast between faces versus houses was used to determine the FFA and STS. The identification of hMT+ was based on the contrast between the movie sequences imitating the face distortion aftereffect versus the first two seconds after the switch from the normal to the contracted face.

Adapation Experiment

Stimuli

Grayscale photographs of faces from the neutral-expression face set of Matsumoto and Ekman (1988) were presented in the centre of the screen. The face stimuli subtended 10° (width) x 13° (length) of visual angle. To investigate the effect of adapting to a distorted face on the appearance of the original version of the face, the photographs were distorted by expanding or contracting the frontal-view image of the face relative to a midpoint on the nose. The distortion was weighted by a circular Gaussian envelope in order to generate the largest changes near the midpoint and only small changes in the outline of the head[1]. Faces of two different identities (face A and B) were distorted.

Procedure

An even block-design was used with 17x18 sec cycles that alternated between a contracted face and a normal face of the same or different identity (Fig. 2c). Five different conditions were included in the experimental design varying in the combination of normal and contracted images (Table 1). Either the contracted face was followed by the original face of the same identity or a different identity. Finally, two control conditions were implemented in the experimental design. During the first control condition the original faces (without distortions) of both identities were presented and in the second control condition an alternation between a uniform and contracted grid was displayed. The order of the conditions was randomized for each subject. Subjects were instructed to view the face stimuli passively.

Table 1
Experimental conditions

Image 1	*Image 2*
Contracted face / identity A	Original face / identity A
Contracted face / identity B	Original face / identity B
Contracted face / identity A	Original face / identity B
Contracted face / identity B	Original face / identity A
Original face / identity A	Original face / identity B
Contracted grid	Original grid

fMRI Data Acquisition

See localizer experiments for general data acquisition details. The adaptation experiments consisted of six different runs (including control conditions), which were presented intermixed with the localizer experiments.

fMRI Data Analysis

See localizer experiments for general data preprocessing details. The data of the adaptation experiments were analyzed using the random effects general linear model approach contrasting the switch from the contracted to the normal face (2 sec; with face distortion aftereffect) versus the switch from the normal to the contracted face (2 sec; no face distortion aftereffect) using a false discovery rate of $q(FDR)<0.01$. The same contrast was used for a detailed analysis within the defined ROIs. Both analyses included all four adaptation conditions.

Results & Discussion

We focused our study on three primary regions of interest (ROIs): (1) the Fusiform Face Area (FFA) known for its responsiveness to faces (Kanwisher, McDermott et al. 1997; Haxby, Hoffman et al. 2000) ; (2) the Superior Temporal Sulcus (STS), a region linked to changeable aspects within faces, such as direction of gaze and biologically relevant face movements (Allison, Puce et al. 2000; Hoffman and Haxby 2000), and (3) the Medial Temporal Area (hMT+), an area sensitive to motion rather than faces. Areas FFA and STS were localized (see Method section) within each subject in a preceding experiment by contrasting face vs. house stimuli. Adaptation scans were acquired while subjects

passively viewed static grayscale face images that alternated at 18 sec intervals between an original face and a face configurally distorted by locally contracting the features toward a midpoint on the nose. Two male faces from the Matsumoto and Ekman (1988) set were used as the normal images. Presentation blocks included normal-distorted alternations between the same or different faces and between the two original faces. Detailed description of procedures are given in the Materials and Method section.

By alternating between the two faces independent of their identity, each face in the pair served both as the adapting and test image. This allowed us to directly compare the magnitude of the aftereffects as the images switched from normal to distorted or vice versa. For this we contrasted responses averaged over a 2 sec aftereffect epoch following each transition (Fig 2c). Patterns of activity for the left and right hemispheres were similar, and thus comparisons of the adaptation aftereffects were based on the pooled responses across both hemispheres and subjects. Figure 2 illustrates both the activation maps (a) and the corresponding average time courses (b). The activation time courses display the fMRI signal during the exposure to different faces (red = original face, yellow = distorted face). For our data analysis only the aftereffect epoch of switching to the original or distorted face image after exposure to the opposite face is considered, However, the reduction of the fMRI signal during prolonged exposure to a given face is also evident at the tail of each activation time course. The responses during the aftereffect epoch in face selective areas were significantly stronger when the faces switched from the distorted to the normal face in both the FFA ($t = 2.73$ $p < 0.01$) and STS ($t = 2.74$, $p < 0.01$). As STS is not sensitive to changes of spatial relations within a face or changes of identity (Rhodes and Michie et al. 2009), the activation patterns are aftereffect related. The observed FFA and STS asymmetry is consistent with a stronger aftereffect in the normal face when subjects are adapted to the distorted face, an asymmetry that is also observed in behavioral measurements of the subjective magnitude of the aftereffects (Webster and MacLin 1999; Leopold, O'Toole et al. 2001; Rhodes, Jeffery et al. 2003; Watson and Clifford 2003; Webster, Kaping et al. 2004; Rhodes and Jeffery 2006). Similarly, all subjects in the present study reliably reported experiencing face aftereffects in the testing conditions, though we did not include a behavioral measure of the aftereffect magnitude. In contrast to the switch between an original and distorted face, alternations between two original faces did not lead to significantly different responses (FFA: $t = 0.43$, $p = 0.66$; STS: $t = -0.06$, $p = 0.95$). This would be expected if the original faces have a similar level of distinctiveness compared to either original and its distortion. In a further control,

significant asymmetries were also absent when observers viewed an alternation between a uniform and contracted grid (FFA: $t = 0.17$, $p = 0.86$; STS: $t = 1.98$, $p = 0.07$), indicating that the observed effects did not depend on the pattern of distortion per se but rather on the distortions in faces.

If human faces are encoded as deviations from a norm face in a continuous space (Leopold et al. 2001) the dissipation of the adaptation induced in our experiment should continuously change the perceived space as the adaptation effect dissipates with time. This is indeed the case for subjects in the contracted face adapted state, exposed to a normal face, reporting the percept of an expanded face trailed by contracting facial movement toward the pre-adapted normal face. To investigate this subjective motion aftereffect we compared responses in the third ROI corresponding to the human Medial Temporal Area (hMT+) (Fig 2a, b). This area is sensitive to motion rather than faces, and was localized by contrasting responses to a single static face versus a movie sequence, mimicking the perceptual post-adaptation aftereffect. The magnitude of response changes in this ROI during the face alternation aftereffect epoch were more similar for the two transitions, though they remained significantly stronger during the aftereffect in the original face ($t = 2.56$, $p < 0.05$), while not differing when switching between two normal faces ($t = 0.61$, $p = 0.58$) or the control grid ($t = 1.36$, $p = 0.19$). Alternation between the normal and contracted face could induce apparent motion (an expansion and contraction) in the facial features, but this motion should be symmetrical for the two transitions. The asymmetry thus implicates a component of the response driven by the perceptual aftereffect of the configural face change. Prior imaging studies have shown activation of hMT+ for passive viewing of illusory motion or static images of implied motion, and during the motion aftereffect when static images are viewed after exposing observers to moving images (Zeki, Watson et al. 1993; Tootell, Reppas et al. 1995; Kourtzi and Kanwisher 2000; Kim and Blake 2007). Our data suggest that hMT+ is also activated by a motion illusion caused by the dynamic perceptual changes during the face aftereffect, as the induced configural changes in the original face fade during renormalization. Notably this component of the motion aftereffect may thus be driven by perceptual changes in the absence of any stimulus motion or implied motion.

Previous tests for a norm in the neural coding of faces have found fMRI responses in the human fusiform gyrus (Loffler, Yourganov, Wilkinson & Wilson, 2005) as well as single cell responses in the inferotemporal cortex of rhesus monkeys (Leopold, Bondar et al.

2006) to be weaker for average than for distinctive faces. These studies have thus provided evidence for the special status of norms in the neural response. Our study extends these findings by instead directly comparing the neural responses following perceptual adaptation to faces. In this case the aftereffects are stronger after adapting to the distorted face, and we show that under these conditions there are correspondingly stronger responses to more neutral faces in the fMRI signal. This similarly suggests that the original face corresponds more closely to a norm, and reveals the neural changes that underlie the recalibration of this norm as observers are exposed to different faces.

The observed asymmetry in the fMRI signal during the face distortion aftereffect epoch is the first evidence of neural responses coupled to illusory post-adaptive face percepts. As noted this asymmetry supports a norm-based code in both FFA and STS and suggests that sensitivity changes at these sites might underlie the face distortion aftereffect and play a role in representing individual variations in faces (though a previous study of aftereffects in the FFA failed to find adaptation to facial categories such as gender or ethnicity; Ng, Ciaramitaro et al. 2006). Our results also point to the critical role that adaptation plays in regulating face norms. To the extent that perceptually more neutral faces are encoded by more neutral response states, adapting to them does not alter face coding since they merely reinforce the existing norm. Conversely, exposure to abnormal faces leads to a recalibration of the norm and consequently to a new pattern of responses – one in which the norm is shifted towards the face the observer is currently exposed to. Coding faces relative to a norm has been contrasted to the visual representation of other dimensions which might instead involve an exemplar code. However, the pattern of face aftereffects and the contrast-coding schemes they imply show many similarities to color vision – in which stimuli are represented relative to a neutral gray, which does not adapt other colors yet is strongly biased by them (Webster, Werner et al. 2005). Thus norm-based codes might in fact represent a general and widespread strategy in perceptual coding.

Adaptation to faces results not only in spatial changes to the perceived configurations but also temporal changes as these configurations renormalize. This creates a powerful motion illusion as the normal test face after adaptation is first perceived as distorted and then morphs over time into the percept of a properly proportioned normal face. Here we show that this motion illusion activates motion coding areas. The inputs driving the neural response to the motion in the aftereffect is uncertain, but to the extent that the asymmetries in the configural aftereffects arise in face selective areas, our results

suggest that these can in turn influence responses to motion in non face-selective areas even there is no direct or implied motion in the stimulus.

In summary our data provide evidence for a norm-based code in human face perception. Systematic adaptation of this coding system creates illusory motion aftereffects powerful enough to activate cortical areas traditionally though to require physical or at least implied motion stimulation. This indicates a hitherto unappreciated abstraction of motion representation in these areas.

References

Allison, T., A. Puce, et al. (2000). "Social perception from visual cues: role of the STS region." Trends Cogn Sci **4**(7): 267-278.

Haxby, J. V., E. A. Hoffman, et al. (2000). "The distributed human neural system for face perception." Trends Cogn Sci **4**(6): 223-233.

Hoffman, E. A. and J. V. Haxby (2000). "Distinct representations of eye gaze and identity in the distributed human neural system for face perception." Nat Neurosci **3**(1): 80-4.

Ishai, A., Schmidt, C.F., Boesiger, P. (2005). "Face perception is mediated by a distributed cortical network.“ Brain Research Bulletin **67**: 87-93.

Kanwisher, N., J. McDermott, et al. (1997). "The fusiform face area: a module in human extrastriate cortex specialized for face perception." J Neurosci **17**(11): 4302-11.

Kim, C. Y. and R. Blake (2007). "Brain activity accompanying perception of implied motion in abstract paintings." Spatial vision **20**(6): 545-60.

Kourtzi, Z. and N. Kanwisher (2000). "Activation in human MT/MST by static images with implied motion." J. Cognitive Neuroscience **12**(1): 48-55.

Kovács, G., M. Zimmer, et al. (2006). "Electrophysiological correlates of visual adaptation to faces and body parts in humans." Neuron **16**(5): 742-53.

Leopold, D. A., I. V. Bondar, et al. (2006). "Norm-based face encoding by single neurons in the monkey inferotemporal cortex." Nature **442**(7102): 572-5.

Leopold, D. A., A. J. O'Toole, et al. (2001). "Prototype-referenced shape encoding revealed by high-level aftereffects." Nat Neurosci **4**(1): 89-94.

Leopold, D. A., G. Rhodes, et al. (2005). "The dynamics of visual adaptation to faces." Proc Biol Sci **272**(1566): 897-904.

Loffler, G., Yourganov G, et al. (2005). "fMRI evidence for the neural representation of faces " Nat Neurosci **8**(10): 1386-1390.

Nielsen, A. F. (2003). "The Brede database: a small database for functional neuroimaging " NeuroImage 19: Presented at the 9th International Conference on Functional Mapping of the Human Brain, 2003 New York, NY

Ng, M., V. M. Ciaramitaro, et al. (2006). "Selectivity for the configural cues that identify the gender, ethnicity, and identity of faces in human cortex." Proc Natl Acad Sci U S A **103**(51): 19552-7.

Rhodes, G. and L. Jeffery (2006). "Adaptive norm-based coding of facial identity." Vision Res **46**(18): 2977-87.

Rhodes, G., L. Jeffery, et al. (2007). "The timecourse of higher-level face aftereffects." Vision Research **47**(17): 2291-6.

Rhodes, G., L. Jeffery, et al. (2003). "Fitting the mind to the world: face adaptation and attractiveness aftereffects." Psychological science : a journal of the American Psychological Society / APS **14**(6): 558-66.

Rhodes, G., P. T. Michie, et al. (2009). "The fusiform face area and occipital face area show sensitivity to spatial relations in faces." Euro J Neurosci **30**(4): 721-33.

Robbins, R., McKone, E., Edwards, M.(2007). “Aftereffects for face attributes with different natural variability: adapter position effects and neural models”. J Exp Psychol Hum Percept Perform. **33**(3):570-92

Talairach, J. and P. Tournoux (1988). "Co-planar stereotaxic atlas of the human brain." New York: Thieme.

Tootell, R. B., J. B. Reppas, et al. (1995). "Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging." Nature **375**(6527): 139-41.

Watson, T. L. and C. W. Clifford (2003). "Pulling faces: an investigation of the face-distortion aftereffect." Perception **32**(9): 1109-16.

Webster, M. A., D. Kaping, et al. (2004). "Adaptation to natural facial categories." Nature **428**(6982): 557-61.

Webster, M. A. and O. H. MacLin (1999). "Figural aftereffects in the perception of faces." Psychon Bull Rev **6**(4): 647-53.

Webster, M. A., J. Werner, et al. (2005). "Adaptation and the phenomenology of perception." Fitting the mind to the world: Adaptation and after-effects in high-level vision **2**: 241–280.

Zeki, S., J. D. Watson, et al. (1993). "Going beyond the information given: the relation of illusory visual motion to brain activity." Proc Biol Sci **252**(1335): 215-22.

Zhao, L. and C. Chubb (2001). "The size-tuning of the face-distortion after-effect." Vision Research **41**(23): 2979-94.

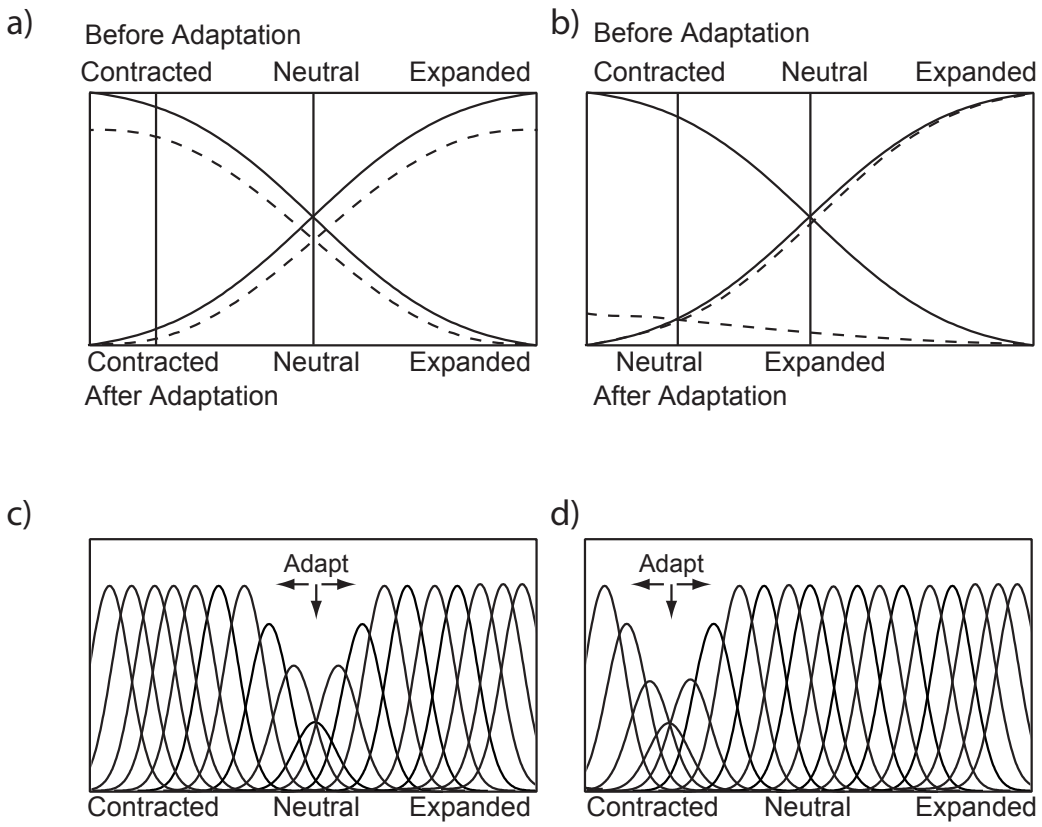


Figure 1

(a & b) The schematic response curves for norm-based face coding, in which a pair of neural mechanisms code deviations relative to a neutral prototype. Solid lines represent the baseline responses, with the normal face corresponding to the cross point at which the two mechanism responses are equal. Adaptation to this normal face does not alter the cross point and thus does not change the face that appears normal or the face that appears contracted (a). Adaptation to a contracted face reduces responses in the mechanism coding contractions. This changes the face that appears neutral and produces a negative aftereffect in the normal face (b). (c and d) Exemplar model, in which the dimension is instead coded by multiple channels. Adaptation reduces responses in the subset of channels tuned to the adapting level. This does not change the modal response or appearance for the adapting level and biases the response to other levels so that they appear less like the adapting level. This predicts symmetric aftereffects after adapting to the neutral face (c) or contracted face (d).

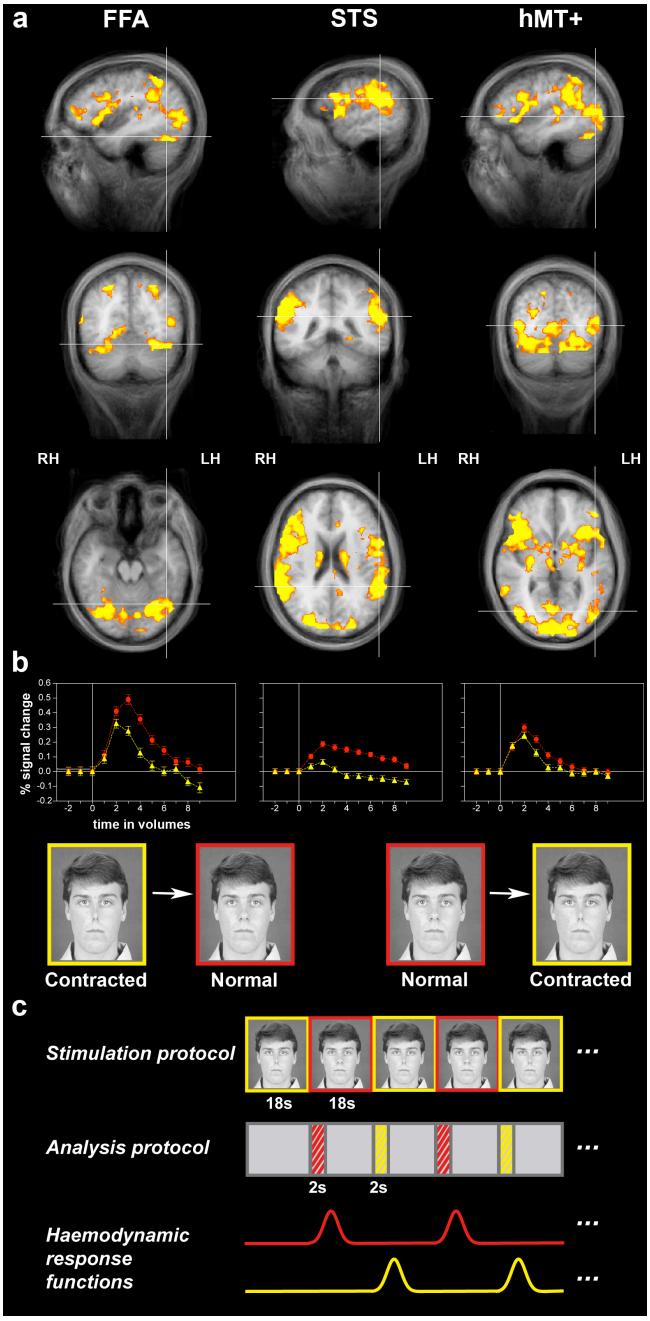


Figure 2

(a) Group activations (yellow-red) contrasting the switch (first 2 sec) of the contracted to the normal face versus the switch of the normal to the contracted face (aftereffect epoch). Activation maps are overlaid on an averaged anatomical T1-weighted dataset (q(FDR)<0.01). Increased activity was obtained in all three ROIs (FFA, STS, hMT+) for the aftereffect following a contracted face compared to the one following a normal face. FFA: Fusiform Face Area. STS: Superior Temporal Sulcus. hMT+: human Medial Temporal Area. RH: right hemisphere. LH: left hemisphere. (b) Average time course signals (±SEM) for each of the three ROIs. Red lines indicate the signal changes following the adaptation to a contracted face, whereas yellow lines represent the signal changes following the adaptation to a normal face. (c) Stimulation protocol. Each face was presented for 18s, alternating between the contracted and the normal face. Analysis protocol and haemodynamic response functions. The first two seconds after each switch between the two faces were analyzed resulting in a reference time course of two predictors (switch from contracted to normal face and switch from normal to contracted face) modeled by the corresponding haemodynamic response functions.

Chapter III

Attention

"(...) as is natural, that of two sensory stimuli the stronger always tends to extrude the weaker, is it conceivable or not that one should be able to discern two objects simultaneously in the same individual time?"

Aristotle
(On Sense and the Sensible)

James (1890) described attention as " the *taking possession by the mind, in a clear and vivid form, of one out of what seem several simultaneously possible objects or train of thoughts".* While attending to one part of our visual field we nearly always move our eyes there, so that the place or object of interest is represented on the fovea; the region of clearest vision. But we can consciously direct our attention away from the point of gaze to attend any place or object within our visual field without moving our eyes. Perhaps we never had conscious awareness of the place or object before we moved our attention there; how is the newly attended place / object represented in relation to the processing of the unattended rest of the visual field?

Before the main part of this chapter will discuss attention as a factor influencing visual sensory processing hence visual perception a more detailed introduction to visual motion processing will be given. As visual areas dedicated to the processing of motion have been described for both monkey (Van Essen, Maunsell and Bixby, 1981) and human (Zeki, Watson, Lueck, Friston, Kennard C and Frackowiak, 1991 ; Watson, Myers, Frackowiak, Hajnal, Woods, Mazziotta, Shipp and Zeki, 1993); the potential to study the influences of attention within these well defined processing regions making use of the well parametrized stimulus space have quickly been realized (for reviews: Wurtz, Goldberg and Robinson 1980; Treue, 2001; Maunsell and McAdams, 2001).

III.I Visual motion processing

Motion perception as a unique sensory process linked to a specific brain circuit was first described by Exner (1894). Moreover, he pointed out moving objects are not captured in a temporal processing of the entire visual field (comparing frame by frame) but rather mark a unique sensation. Today we know that the primate brain not only contains specialized neurons processing visual motion but that entire cortical areas are devoted to the processing of motion.

III.I.I Visual areas involved in motion processing

The first brain region known to process motion is V1; Hubel and Wiesel (1968) described extracellularly recorded neurons in the macaque brain to be more responsive to one direction of motion while showing little or no response when the stimulus was moved in the opposite direction. Within the functionally distinct six layers of V1 neurons in the upper layer 4 (4A, 4B, 4Cα) (with 4Cα receiving the most magnocellular input from the lateral geniculate nucleus (Hubel and Wiesel, 1972)) and neurons in layer 6 display direction selective responses (Hawken, Parker and Lund,1988). Projections onward from layer 6 (Fries, Keizer and Kuypers, 1985) and 4B (Lund, Lund, Hendrickson, Bunt and Fuchs, 1976; Maunsell and VanEssen,1983) of the primary visual cortex have been directly linked to the middle temporal area (MT) within posterior bank of the superior temporal sulcus (STS) (Zeki, 1974; Maunsell and Van Essen,1983). Additional to the direct input of V1 neurons afferent connections of other visual areas, namely V2, V3 to area MT have been identified.

Like V1, MT has a columnar organization, neurons preferring similar directions are clustered next to one another (Albright, Desimone and Gross, 1984; DeAngelis and Newsome, 1999). The majority of MT neurons are direction selective (Albright, Desimone and Gross, 1984) with single neurons representing the stimulus direction precise enough to account for the behavior in a motion direction discrimination task (Britten, Shadlen, Newsome and Movshon, 1992). Receptive fields (RFs) of MT

neurons, restricted to the contralateral hemifield are ~ 10 times larger then V1 RFs, with their diameter related to their eccentricity (factor of 0.8) (Maunsell and Van Essen, 1983). VanEssen, Maunsell and Bixby (1981) were the first to describe strong direction selective cells located medially on the upper bank of the sulcus adjacent to area MT with exceptionally large excitatory receptive fields. These receptive fields were about 14º larger than those of MT at all eccentricities (Desimone and Ungerleider, 1986). This area, later defined as medial superior temporal area (MST) (Maunsell and VanEssen, 1983) receives its primary input (with returning feedback connections) directly from area MT (Ungerleider and Desimone, 1986, Boussaoud, Ungerleider and Desimone,1990). Additionally, ascending connections between areas V1, V2, V3 onto MST have been identified; MST itself has afferent fiber projections into the ventral and dorsal pathway namely: FST (fundal surface of the STS; engaged in motion segmentation; Rosenberg, Wallisch and Bradley, 2008), LIP (lateral intraparietal area; saccade planning area / visuospatial attention; Colby and Goldberg, 1999), PO (parieto-occipital area; Visuomotor related responses; Galletti, Battaglini and Fattori, 1995), DP (dorsal prelunate area; activated by eye movements and optic flow; Li, Tanaka and Creutzfeldt , 1989; Raffi and Siegel, 2007), area TEO (Shape detection and processing; Boussaoud, Desimone and Ungerleider,1991) and the frontal eye field located in the prefrontal cortex (Barbas and Mesulam, 1981; Boussaoud, Ungerleider and Desimone, 1990).

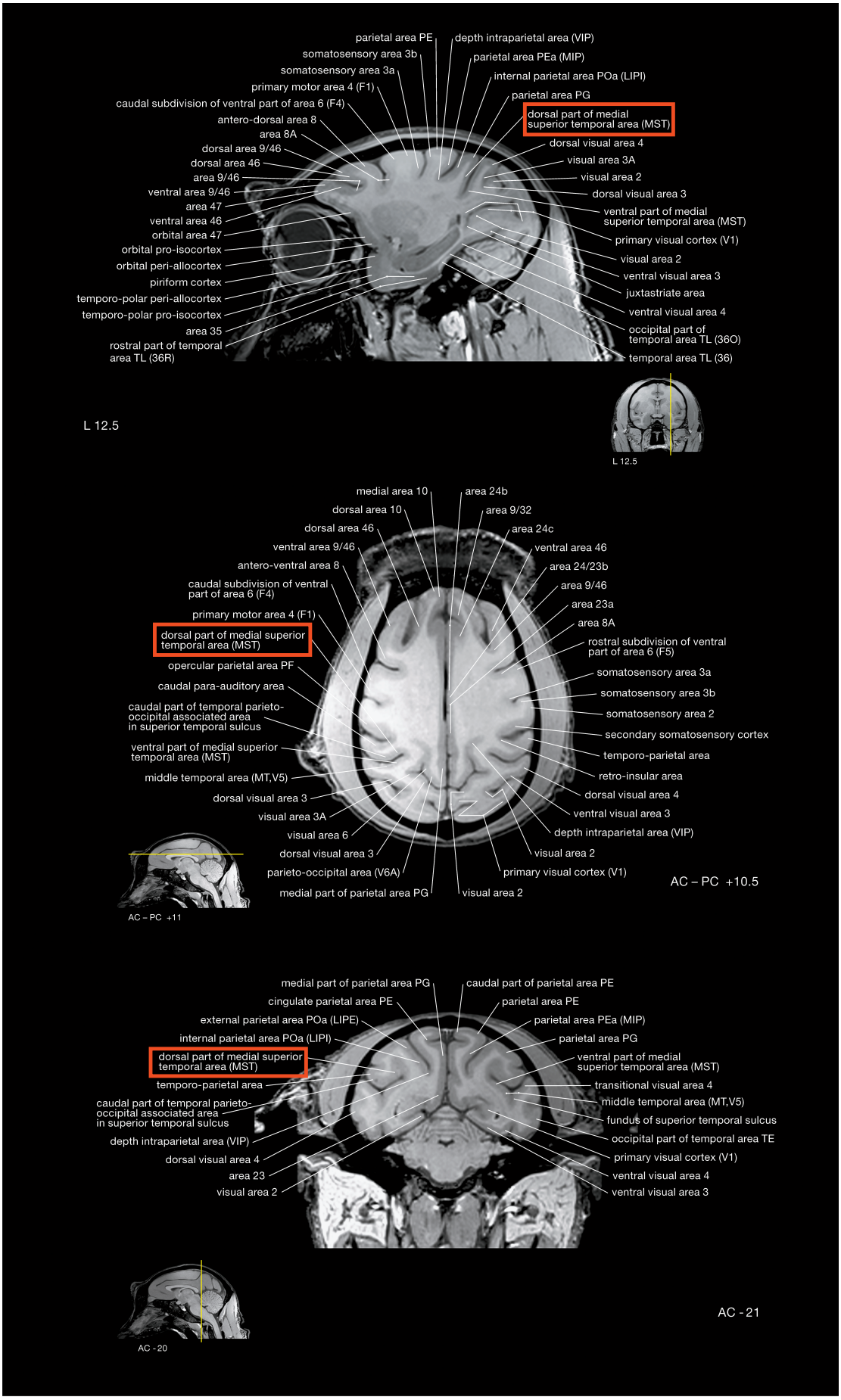
There is evidence for a crude visual topography within area MST (Desimone and Ungerleider, 1986) along some myeloarchitectural differences. A densely myelinated zone in the upper bank of the macaque STS, contains cells with predominantly peripheral visual fields and a lightly myelinated zone located posterolateral of the densely myelinated area with cells having predominantly central visual fields. Due to the architectural structure of MST it was suggested to contain more than one functional purpose (Desimone and Ungerleider, 1986; Komatsu and Wurtz, 1988).

III.I.II Functional properties of area MST and the perception of motion

The motion sensitive areas in the medial superior temporal area have been divided into the dorsal (MSTd) and the lateral (MSTl) region. MSTl neurons, described as tracking neurons, reduce textured background movement from the retina to allow for a world-centered representation of a target. These neurons, activated during a target pursuit remained active even when the tracked visual stimulus is turned off (Ilg, Schumann and Thier, 2004) and the eyes follow along the trajectory of the stimulus to an anticipated endpoint. Further support of the tracking neuron characteristic is observed by the pronounced interference in eye movements following lesions to MSTl (Dursteler and Wurz, 1988).

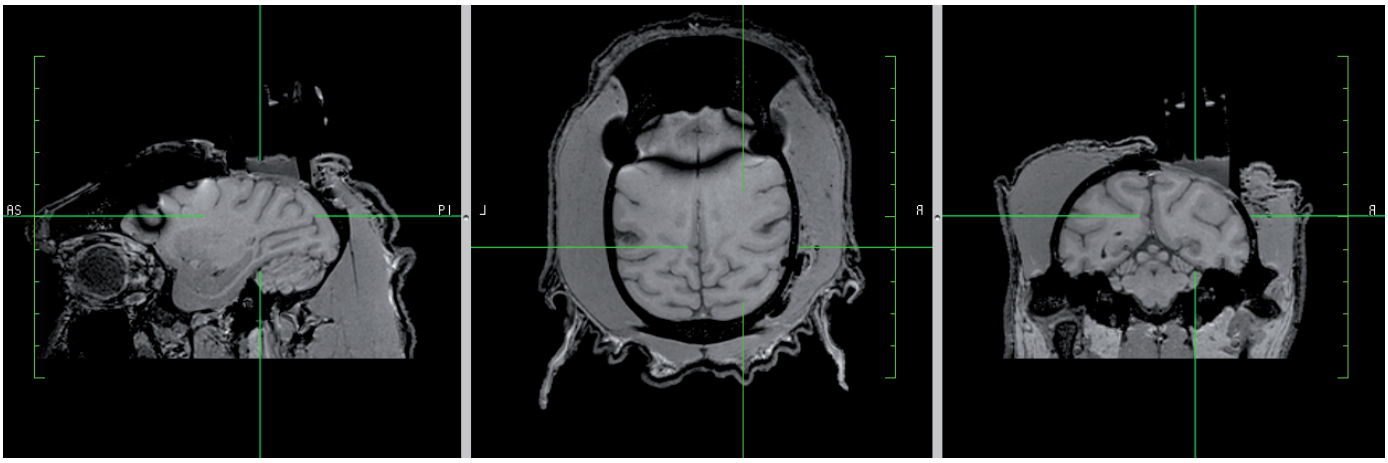
Cells in the dorsal region (MSTd) prefer more complex, global motion (optic flow) patterns simulating experiences of self motion (Heuer and Britten, 2004). Their large excitatory receptive fields extend not only across the contralateral hemifield but additionally into the ipsilateral visual field often including the foveal region (Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986; Tanaka and Saito, 1989). Though successively recorded cells share similar RF size and location (Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986) no correlation between eccentricity and size has been found. At first MSTd cell responses were crudely determined using circular rotational and expanding / contracting (radial) motion patterns (Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986; Tanaka and Saito, 1989; Duffy and Wurtz, 1991) moving at the preferred speed of the neuron. A well parameterized spiral motion space (SMS), evoking reliable single cell response differences was established by Graziano, Andersen and Snowden (1994). This space can best be described as a coordinate system consisting of two orthogonal axes, one representing clockwise / counterclockwise rotation the other expansion / contraction. Stimuli of intermediate directions are a combination of expanding or contracting plus a rotational (clockwise / counterclockwise) stimulus component to create a smooth transitions between the four cardinal directions.

The direction selectivity of single MSTd neuron can be characterized by the SMS tuning profile usually well-fitted by a Gauss function. That is, an individual MSTd cell will give its strongest response to a particular SMS motion pattern (e.g. clockwise expanding spiral - preferred direction (Pref Dir)) while a stimulus moving in the opposite direction (counterclockwise contracting spiral - anti-preferred direction (AntiPref Dir)) will elicit little or no response. All other stimuli within the SMS will evoke intermediate responses, their magnitude depending upon their similarity to the Pref Dir/ AntiPref Dir. The response of a single MSTd cell during a discrimination task varying the coherence of motion can account for the animals choice, as the cells threshold parallels / exceeds the behavioral choice (Heuer and Britten, 2004). A hallmark of MSTd is clearly its wide range of sensitivities. Many cells not only respond to SMS stimuli but in addition exhibiting a response profile similar to MT when exposed to linear motion stimuli (Tanaka, Hikosaka, Saito, Yukie, Fukada, Iwai, 1986; Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986; Graziano, Andersen and Snowden, 1994). Further, MSTd responses remain selective over the entire RF; cells directionally selective responses are preserved irrespective of stimulus size and location (Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986; Duffy and Wurtz, 1991; Graziano, Andersen and Snowden, 1994; Geesaman and Andersen, 1996).

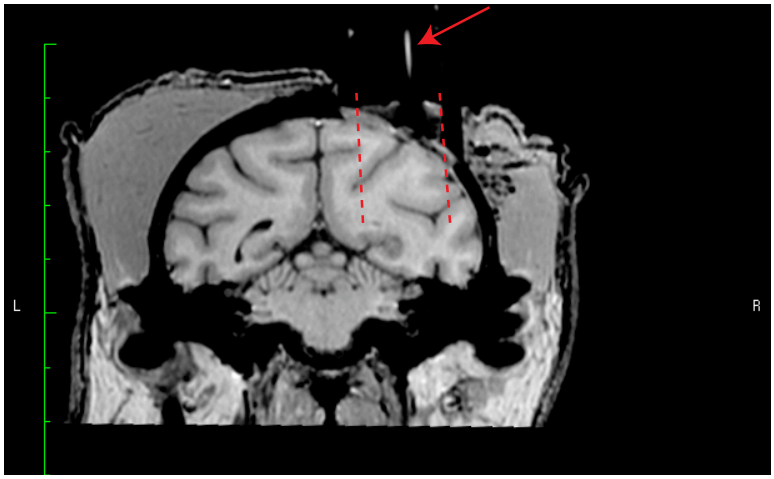


Magnetic Resonance macaque brain scans; top to bottom (sagittal plane, transverse plane, coronal plane) showing reference images identifying the dorsal region of the medial superior temporal (area MSTd - red box) (from Tammer, Hofer, Merboldt and Frahm, 2009)

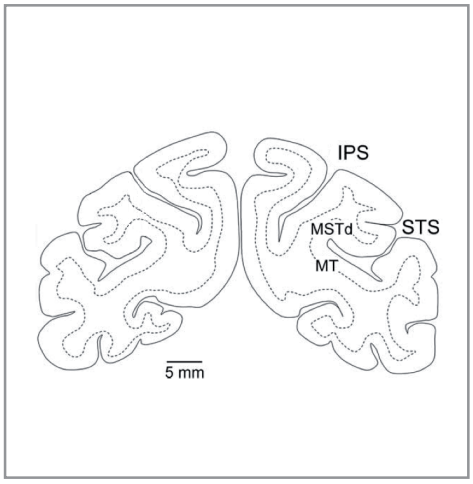
a)



b)



c)



Recording location of one macaque monkey participating in the electrophysiological attention studies presented in this thesis. The Magnetic resonance images in the top row (a) show sagittal, transverse and coronal slices of the recording chamber giving access to the superior temporal sulcus with the MSTd recording site (green cross). Picture (b) shows a close-up of the coronal plane; the red arrow show a water filled capillary inside the recoding chamber (red dotted line) indicating the orientation of the electrode relative to the brain surface. Schematic representation of area MSTd is depicted in image (c).

III.II Attention - response modulation

Early neurophysiological studies reported behavior related enhanced neuronal responses in macaque monkeys while these were engaged in saccadic eye movement task (Goldberg and Wurtz, 1972; Mountcastle, 1976; Yin and Mountcastle (1978 Wurtz, Goldberg and Robinson 1980). The firing rate of the neuron increased when the behaviorally relevant stimulus (positioned within the receptive field) became the relevant saccade target. The modulation of motion dedicated cortical neurons (in area MT and MST) during the tracking of a visual moving target (Newsome, Wurtz and Komatsu, 1988) did not allow for a decisive isolation of attentional effects. As modulatory eye movement changes evoked by the stimulation of MT and MST neurons (Komatsu and Wurtz,1989) identify that stimulus related eye movement initiation itself increases neuronal responses.

More recent studies forwent the constricting interaction of attention and eye movements. Establishing behavioral tasks of allocating attention in the absence of planing or executing directional eye movements toward a motion stimulus. While holding eye fixation constant macaque monkeys were trained to attend to one peripheral random dot pattern (RDP) target stimulus in the presence of another RDP (distractor stimulus). One of the RDPs was placed in the RF while the other was placed outside, usually in the opposite hemifield. The animal reports a change (speed or direction) in the behaviorally relevant RDP stimulus either placed inside or outside the RF in the presence of the distractor stimulus moving in the same direction as the target stimulus. Neurons in motion sensitive area MT respond more strongly when the stimulus inside the RF is attended compared to when the stimulus is unattended (attention directed to the identical RDP outside the RF (Treue and Maunsell, 1996)). This location dependent sensory information enhancement (spatial attention) acts along all stimulus dimensions that the neuron is selective for.

Against earlier research suggesting that spatial attention changes the selectivity of a neuron (Spitzer, Desimone and Moran, 1988) more recent studies clarified that attention

multiplicatively enhances responses along the entire tuning profile (Treue and Martinez-Trujillo,1999; McAdams and Maunsell, 1999). Consequently, the responses to all stimuli along the entire tuning curve are proportionally scaled by one factor (response gain model; McAdams and Maunsell, 1999 Williford and Mausell, 2004), without broadening or sharpening tuning curve width (i.e. selectivity of a cell). Further, directing spatial attention to one of two motion stimuli inside the RF reliably shift the center of the RF (Womelsdorf, Anton-Erxleben, Pieper and Treue, 2006) along with the inhibitory surround (Anton-Erxleben, Stephan and Treue, 2009) toward the attended target. This might contribute to the spatial segregation of a relevant stimulus out of a cluttered surround.

Modulatory attentional effects can also be observed independent of the attended location within the visual field. Attending to a particular feature (Feature-based attention e.g. upward motion, color red) enhances the response of all neurons participating in the processing of that particular feature. Treue and Martinez-Trujillo (1999) showed that attending to the Pref Dir (motion feature) of a neuron at a distant (outside) RF enhances the response of the unattended Pref Dir stimulus inside the RF, while attending to the AntiPref Dir outside reduces the response. The feature-based attention motion tuning curve alike the spatial attention motion tuning curve did not vary in width or sensitivity. However it reflected the similarity of the attended feature (motion direction) to the preferred feature of the recorded neuron (feature-similarity gain; for review: Maunsell and Treue, 2006).

III.II.I Attention - progression & synchronization

Progressing through the processing stream indicates that the magnitude of the attentional modulation increases (Treue, 2001; Maunsell, 2003). Single cell response intensity show weak increases to attended stimuli in striate visual cortex (V1) that become stronger in extrastriate cortex (i.e. V4, MT)(Luck, Chelazzi, Hillyard and Desimone, 1997; Treue and Maunsell, 1999; McAdams and Maunsell, 1999; Seidemann and Newsome, 1999). The bottom-up visual signal and the top-down attentional mechanism must interact to amplify and transform the relevant information to

create a saliency map accentuating regions of interest in the visual field based upon stimulus strength and behavioral relevance (Treue, 2003). This suggests some form of coupling of selected groups of neurons that communicate information about the stimulus and its behavioral relevance not only locally but also on a long-range scale. While firing rates of a single neuron are highly variable on a trial by trial basis, even when a physically identical stimulus is shown (Shadlen and Newsome, 1998) attention selectively enhances not only the single neuron responses, it also synchronizes neuronal population activity (Womelsdorf and Fries, 2007). Fries, Reynolds, Rorie and Desimone (2001) measured local field potentials (LFP) (low pass filtered extracellular recording signals) and the spike train of a V4 neuron (high pass filtered) while a macaque monkey engaged in a spatial attention task. When the animal attended a stimulus inside the RF the coherence in the gamma frequency (40–100 Hz) range of the LFP and the spike train increased (for review: Fries, 2005). Therefore, this synchronization may be a general mechanism for regulating communication across cortical areas.

III.III MSTd and attention: a short outline

Throughout the visual motion processing hierarchy neurons can best be described as filters, coding information selectively along specific stimulus dimensions (VanEssen, Anderson and Fellman, 1992). Processing efficiency can be enhanced for stimuli in the focus of attention resulting in faster detection times, lower perceptual thresholds and higher judgement of accuracy (Posner, Snyder and Davidson, 1980). Thus, the active process of perception does not merely reflect the bundling of extraretinal stimulus information overlapping the neurons receptive field. It can be described as the modulation of the visual information according to its behavioral relevance.

Attention enhances the responses of neurons across the visual cortex, when directed to a behaviorally relevant location within the visual field (spatial attention) or a particular feature (feature-based attention). The most systematic spatial attention studies have investigated attentional effects onto the tuning curve to the most relevant stimulus dimension for a given neuronal population. In the following manuscripts this approach will be used to test if MSTd tuning profiles of individual neurons and therefore the response of populations of MSTd neurons can be well fitted by a Gauss-function. These

fitted tuning curves allow for a qualitative comparison of neuronal responses for behaviorally relevant (attended) or unrelevant (unattended) stimuli at different spatial positions.

Allocating attention to a feature creates a location-independent modulation of responses across the entire visual field. The feature-similarity gain model predicts attentional modulation of MSTd neurons to reflect the similarity of the behaviorally relevant feature and the sensory selectivity of the neuron along this dimension. The magnitude of the feature-based attentional modulation will be subject matter in the following part. The spatial and feature-based attentional modulation effects within area MSTd will be compared to the responses of prejacent motion processing area MT.

Attention - manuscripts

- Kaping D, Baloni S, Zhang L and Treue S. Spatial attention modulated MSTd responses to optimized spiral motion space stimuli. Prepared for submission

  Author contribution: DK and ST designed the original experiment; DK and SB implemented the experiment, collected and analyzed data. LZ analyzed the waveform. DK wrote the main paper, ST edited the manuscript; all authors discussed the results and commented on the manuscript at all stages.

- Kaping D, Baloni S and Treue S.. Feature-Based Attention in MSTd; not all stimuli giving rise to a response are treated as a feature. Prepared for submission

  Author contribution: DK and ST designed the original experiment; DK and SB implemented the experiment, collected and analyzed data; DK wrote the main paper, ST edited the manuscript; all authors discussed the results and commented on the manuscript at all stages.

III.III.I Spatial attention modulates activity of single neurons in primate visual cortex.

The allocation of spatial attention modulates neural responses to visual stimuli in many areas of primate visual cortex. The most systematic studies have investigated attentional effects onto the tuning curve to the most relevant stimulus dimension for a given neuronal population such as orientation tuning for V4 neurons or tuning for the direction of linear motion for MT neurons. Here we use this approach for MSTd neurons which have a characteristic and prominent tuning for direction in spiral motion space (SMS). SMS has been introduced by Graziano et al. (1994) as a circular dimension that considers expansion, clockwise rotation, contraction and counterclockwise rotation as neighboring stimuli in this space, with a continuum of stimuli in between these cardinal directions.

We recorded SMS responses from MSTd neurons of two macaque monkeys. The monkeys were trained to attend to one spiral motion random dot pattern (RDP) target stimulus in the presence of another RDP (distractor) while maintaing their gaze on a fixation point. One of the RDP was placed in the receptive field (RF) while the other was placed outside, usually in the opposite hemifield. In one of three conditions the monkeys reported a luminance change of the fixation point while the unattended RDP inside and outside the RF moved in one of twelve possible SMS directions. In the remaining two conditions behaviorally relevant target stimuli either inside or outside the RF moved in one of twelve possible SMS directions in the presence of a distractor stimulus. The monkeys reported a speed change within the target stimulus while ignoring all changes within the distractor stimulus. The tuning profile of individual MSTd neurons and therefore the response of populations of MSTd neurons can be well-fitted by a Gauss function. These fitted tuning curves allow for a quantitative comparison of neuronal responses for behaviorally relevant (attended) or unrelevant (unattended) stimuli at different spatial positions.

Spatial attention modulates MSTd responses to optimized spiral motion space stimuli

Daniel Kaping[1,2], Sonia Baloni[1,2], Lu Zhang[1,3], Stefan Treue[1,2]
[1]Cognitive Neuroscience Laboratory, German Primate Center, Goettingen, Germany
[2]Bernstein Center for Computational Neuroscience, Goettingen, Germany
[3]Shanghai Institute of Brain Functional Genomics, East China Normal University, China

Introduction

It is well established that attention modulates neuronal sensory signals in many areas of primate visual cortex. Directing attention to a behavioral relevant target stimulus placed inside the receptive field of a neuron multiplicatively enhances responses along the entire tuning profile (Treue and Martinez-Trujillo,1999; McAdams and Maunsell, 1999). Consequently, response tuning curves are proportionally scaled by one factor (McAdams and Maunsell, 1999; Williford and Mausell, 2006), without broadening or sharpening tuning curve width (i.e. selectivity of a cell). The observed response enhancement differs between cortical areas progressively increasing (Treue, 2001; Maunsell and Cook, 2002) along the visual processing hierarchy (Maunsell, 1995). Weak increase to an attended stimulus are observed in striate visual cortex (recently observed even earlier in Lateral Geniculate Nucleus, McAlonan and Wurtz, 2008) that become stronger in extrastriate cortex (Motter, 1993; Treue and Maunsell, 1996; Luck, Chelazzi, Hillyard and Desimone, 1997; McAdams and Maunsell, 1999; Seidemann and Newsome, 1999; Cook and Maunsell, 2002). Along the dorsal motion processing pathway (Maunsell and Van Essen, 1983) signals in the middle temporal area (MT) are weaker compared to other motion selective areas ascending within the superior temporal sulcus. Progressive attentional enhancements can be observed for delayed match to sample tasks (Ferrera, Rudolph and Maunsell, 1994), target pursuit (Recanzone and Wurtz, 2000) and speed / direction change detection tasks (Treue and Maunsell, 1996). Spatial attention induced response enhancements to unidirectional linear motion moving in the preferred direction of a cell were reported not only for cells in area MT but as well for cells in the dorsal region of medial superior temporal area (MSTd) (Treue and Maunsell, 1996). While cells in MSTd, receive strong direct projections from area MT and respond well to linear motion stimuli they prefer more

complex, global motion patterns (Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986; Tanaka and Saito, 1989; Duffy and Wurtz, 1991) simulating experiences of self motion / optic flow (Roy and Wurtz, 1990). Based on this evidence we ask how responses to optimized MSTd stimuli (complex motion) are modulated by spatial attention. Best suited to systematically study the effect of attention onto MSTd responses is the well parameterized spiral motion space (SMS) (Graziano, Andersen and Snowden, 1994). It can best be described as a circular space that considers expansion, clockwise rotation, contraction and counter clockwise rotation as neighboring stimuli forming a continuum of stimuli in between these cardinal directions. Tuning profiles of individual MSTd neurons and therefore the response of populations of neurons can be obtained by means of reliable response differences to SMS stimuli. The well fitted Gaussian tuning curves allow for a qualitative comparison of behaviorally relevant (attended) and unrelevant (unattended) stimuli at different spatial positions. Based on findings of the profile of attentional modulation in prejacent area MT we anticipate spatial attention to create a multiplicative change to SMS tuning functions in area MSTd. Alternatively cells could become more selective through an attentional narrowing in tuning width. We therefore try to distinguish whether attention creates a purely multiplicative change or combines enhancement and compression along the tuning curve.

Our results document that directing spatial attention into the receptive field of MSTd neurons strengthens their response. The observed gain change to optimized complex motion patterns was 30% on average with response enhancements along the entire tuning profile. Thus, spatial attention modulates responses compared to that of area MT (reported to be roughly 15%) twice as strong supporting the notion of a visual location dependent attentional boost progressing up the visual processing hierarchy. Utilizing known waveform differences of extracellular recorded neurons (Mitchell, Sundberg and Reynolds, 2007; Johnston, DeSouza and Everling, 2009) we found similar attentional effects in pyramidal neurons and inhibitory interneurons within MSTd microcircuitry.

Materials and Methods

We recorded responses of 121 well isolated motion-selective neurons in area MSTd of the visual cortex in two hemispheres of two macaque monkeys (monkey N: n = 65 and monkey W: n = 56) engaged in a spatial attention task. Experiments were performed in a dimly lit room. During the experiment, a custom computer program running on an Apple Macintosh PowerPC controlled the stimulus presentation, monitored and recorded eye positions, neural and behavioral data. Eye positions were determined using video-based eye-tracking (ET49, Thomas Recording, Giessen, Germany). Monkeys sat in a custom-made primate chair viewing visual stimuli on a computer monitor (distance of 57 cm). The monitor covered 40º x 30º of visual angle at a resolution of 40 pixel/deg.

Stimuli

We used random dot patterns (RDPs) of small bright dots (density: 8 dots per degree^2, luminance 75 cd/m^2) plotted within a stationary circular aperture on a gray background of 35 cd/m^2. Stimuli were spiral motion space (SMS) patterns considering expansion, clockwise rotation, contraction and counterclockwise rotation as neighboring stimuli with a continuum of stimuli in between these cardinal directions. The direction of SMS was determined by the angle that the individual dot formed with radial reference lines. By varying the angle we are able to create smooth transitions between neighboring directions within spiral motion space (SMS). Movement of the dots was created by the appropriate displacement of each dot at the monitor refresh rate of 75Hz.

Animal preparation and neural recordings

Following initial training, monkeys were implanted with a custom-made orthopedic implant preventing head movements during training and extracellular recording. A recording chamber was placed on top of a craniotomy over the left (monkey N: 3.25 mm posterior/16.3 mm lateral; Crist Instruments, CILUX Recording Chamber 35º, Hagerstown, MD) or the right (monkey W: 3 mm posterior/ 15.5 mm lateral; custom-fit computer-aided milled magnetic resonance imaging (MRI) compatible chamber, via digitized monkey skull surface reconstruction, 3di, Jena, Germany) parietal lobe. Pre-surgical MRI was used to position the chambers; post-surgical MRIs verified the correct positioning and precise targeting of area MST. All surgeries were performed under general anesthesia and sterile conditions. Animal care and all experimental procedures were conducted in accordance with German laws governing animal care and approved

by the animal ethics committee of the district government of Braunschweig, Lower Saxony, Germany.

For extracellular recordings we simultaneously used up to three microelectrodes in a three-electrode twelve channel system (Mini-Matrix, Thomas Recordings, Giessen, Germany). The dura mater was penetrated with sharp guide tubes so that electrodes could be inserted into the brain. The raw signal of the electrodes was amplified (gain range 1000 - 32 000) and filtered (frequency range 150-5000 Hz). Action potentials were online-sorted (waveform window discrimination, Sort Client, Plexon Inc., Dallas, TX) and recorded.

Receptive fields (RFs) of well-isolated single MSTd cells were identified by responses to a stationary RDP stimulus manually swept across the screen. To characterize the SMS direction and speed of the individual cell the monkey performed a luminance task on a centrally positioned fixation point, while a RDP was presented within the estimated MSTd RF. The size of the RDP was matched to allow the placement of two RDPs at equal eccentricity to the fixation point (inside & outside the RF). Twelve SMS directions at a maximum velocity of 8 degrees per second for the dots furthest away from the center were randomly chosen in intervals of 827 ms . Responses of the individual SMS direction were defined as the mean firing rate in an interval of 80-800 ms after onset of a particular SMS direction. A SMS direction tuning curve was fitted online with a circular Gaussian. The SMS direction yielding the highest mean firing rate was presented at eight different speeds (between 0.5 and 64 deg/sec) to determine the preferred speed of the individual neuron.

Behavioral Tasks

For the main experiment, two monkeys were trained in a spatial attention task (Figure 1) to attend one SMS RDP target stimulus in the presence of another distractor RDP. A single trial started with the presentation of a fixation point (0.2º x 0.2º) placed in the center of a computer screen and a stationary cue RDP (size, dot density) positioned at the location to be attended throughout the trial. Once the animal foveated the fixation point and touched the response lever the cue RDP turned off. While maintaining their gaze on the fixation point a brief blank period (225 ms) was followed by the onset of two simultaneously presented random motion RDPs (375 ms). One RDP was placed in the RF while the other was placed at equal eccentricity to the fixation point outside the RF

(opposite hemifield). In a given trial the two RDPs changed to coherent SMS motion, moving in the same direction, picked randomly from one of twelve directions in SMS. This ensures that feature-based attention is equated between any pair of conditions. The monkeys had to report a speed change within the cued target stimulus while ignoring all changes within the distractor stimulus. The target was either the stimulus inside (attention-in condition) or outside (attention-out condition) the RF. Changes of the target stimulus and distractor stimulus occurred at random times between 150 ms and 2400 ms after the onset of coherent SMS motion (total possible trial length 3000 ms). Our experiment allowed us to compare the neuronal responses during the sustained state of selective spatial attention to the stimulus inside the RF, when behaviorally relevant (attended target stimulus) or irrelevant (unattended distractor).

Data Analysis

Data were analyzed offline with custom scripts using MATLAB (The Math Works, Natick, MA). For the analysis of neuronal data only correctly performed, completed trials were included. Spike density functions (SDF) were obtained by convolving each spike with a Gaussian function (σ = 20). Response rates were determined by averaging the frequency of action potentials over 850 ms of stimulus presentation starting 267 ms after coherent SMS onset within a trial and then pooling across trials of the same SMS directions. SMS tuning curves are generally well represented by a Gaussian function (Graziano, Andersen and Snowden,1994). We fitted SMS tuning functions for both tasks (attention-in and attention-out). The function had four free parameters : a = asymptote, b = amplitude, c = tuning width, d = preferred direction and s = spontaneous firing rate of the neuron.

$$f(x)=s+a+b\cdot e^{\left(-\frac{(x-d)^2}{2c^2}\right)}$$

Population responses were computed averaging across normalized responses in the attention-in condition. The effects of spatial attention were analyzed during the sustained response after the onset of coherent SMS stimuli. During all trials the RDPs (inside and outside the RF) moved in the same direction. For each recorded neuron the differences in response between the attended and unattended condition were computed (AI (attentional modulation index : [(attention-in - attention-out)/(attention-in + attention-

out)])). The average AI corresponds to the geometric mean of the ratios and thus is a conservative measure of the average attentional effect, reducing the effects of outliers. The statistical significance of the AI was evaluated using a one-sample t test.

The average waveform of each neuron was normalized into z-score and then spline interpolated to resolution of 1 s. Then waveforms were aligned by their troughs (average waveform see Figure 2a). The duration of waveforms was computed by the distance between the trough and the peak. The distribution of waveform was significantly bimodal (Hartigan dip test, $p < 0.01$, Figure 2b). Based on the bimodal distribution, narrow-spiking neurons was defined as those duration from 100 to 195 μs and broad-spiking neurons was defined as those duration from 196 to 600 μs .

Results

We examined the effects of attention on the SMS direction tuning of 121 isolated neurons in area MSTd in two macaque monkeys. While there exist conflicting reports on rotation, expansion / contraction selective cells within area MSTd (Lagae, Maes, Raiguel, Xiao and Orban, 1994, Duffy and Wurtz, 1995) Graziano, Andersen and Snowden (1994) reported 86% (57 of 67) of cells tested with expansion, contraction, CW / CCW rotation and four intermediate SMS directions (expanding CW spiral, expanding CCW spiral, contracting CW spiral and CCW contracting spiral) to be well direction selective. As in the study mentioned above we determined the selectivity of MSTd neurons to unattended SMS stimuli. The animals engaged in a luminance change detection task on the fixation spot while one of twelve possible SMS directions was presented inside the RF. Here we plot the SMS resulting tuning curves in Cartesian coordinates (fit with a Gaussian) with the mean (peak of the Gaussian) to correspond to the preferred direction and to provide a measure of width (σ, the SD of the Gaussian) and goodness of fit (r, the correlation coefficient). Repeated presentation of a set of 12 SMS directions to individual recorded MSTd neuron allowed for a discrete conformation of the preferred direction. For the example neuron (Figure 3) clockwise rotation was the preferred SMS direction; error bars show SE (computed across n=10 trials) displaying the confidence window for the mean. The width of its tuning curve was 57.3º with a goodness-of-fit of $r > 0.98$. This response is exceptional, however, not an atypical observation in population. In total, 48 cells (40%) had smooth tuning curves with a good Gaussian fit $r > 0.9$ and 66% above $r = 0.8$. For the population of 121 cells the average

goodness-of-fit was r = 0.78 along with an average width of 61º (σ). The tuning of the 121 MSTd cells to the SMS stimuli is quite pronounced. With the preferred SMS motion response on average more than 4 times higher (directionality = 0.76) then the response to the anti preferred SMS direction (Figure 4a).

As MSTd cells are known to be selective for a range of stimuli (Duffy and Wurtz, 1991) we tested a subset of 61 neurons for their selectivity to unidirectional linear motion. Linear motion tuning curves had good Gaussian fits, over 57% (35 neurons) better then r = 0.9, (over 75% r < 0.8). We observed a significant difference in absolute firing rate between the SMS tuning and the linear motion tuning. That is, MSTd neurons exhibit 33% higher responses to the preferred SMS motion stimulus over the preferred linear motion stimulus (Figure 4b).

Effects of attention

Responses of the preferred SMS direction were strongly modulated when the animal was attending to the stimulus inside the RF compared to attending away. For the analysis of neuronal data only correctly completed trials are included. In all trials the RDP inside the RF moved in one of twelve possible SMS direction the physically identical stimulus was presented outside the RF. The animal was cued to one spatial location inside or outside the RF to report a speed change within the presented SMS pattern. Data, plotting the time course of an example neuron recorded while the animal attended / unattended the preferred SMS RDP pattern is shown in Figure 5. The top part plots spike trains where each train was acquired during a single trial of attending to the preferred SMS direction inside the RF (blue) or attending away from the preferred SMS direction inside the RF to the identical SMS RDP presented outside the RF (red). Response rates were determined by convolving the spike train in each trial with a Gaussian kernel (σ = 20 ms); and averaging across trials. While attention directed to the stimulus placed outside the receptive field (red) evokes a strong response of the neuron to the coherent preferred SMS stimulus and a fairly stable level of activity after the initial transient response; directing attention inside the receptive field triggers a much higher response to the identical sensory input. The response for the attended condition over the sustained response period (gray shaded) was over 50% greater (51 spikes/sec versus 30 spikes/sec). Data of the sustained population response to the preferred SMS direction of 121 neurons show a clear effect of attention (Figure 6). For

each recorded neuron, the attentional index (AI) was computed (Figure 7): AI = ($prefDir_{Attend\ IN}$ – $prefDir_{Attend\ OUT}$)/($prefDir_{Attend\ IN}$ + $prefDir_{Attend\ OUT}$), where prefDir is the mean response to the preferred direction inside the RF once as a target $prefDir_{Attend\ IN}$ once as a distractor $prefDir_{Attend\ OUT}$. Responses were computed by averaging across the normalized responses of the preferred direction showing an average increase of 30%. The significance of the AI was statistically evaluated using a one-sample t test showing a significant attentional effect $p < 0.01$ (AI of 0.13). As the response to the preferred SMS stimulus creates such strong response enhancements the main question to answer remains whether spatial attentional modulates entire MSTd tuning curves in a similar fashion as observed in predjacent area MT. Recording responses of 12 SMS directions once attended, once unattended allowed us to evaluate the change in tuning parameters (Amplitude, Asymptote, Width and Direction) of the 121 MSTd neurons. Neurons with significant effects of attention to the preferred SMS direction had an increase in responses across the entire tuning curve for attend inside the RF condition. The change in amplitude of the SMS tuning curves resulting from the allocation of attention (Figure 8a) are identical to the previous results plotted in Figure 6 and Figure 7; as the amplitude change (peak of the Gaussian) represents the attention induced change of the preferred direction. The mean Amplitude Index of the population is shifted to the right 0.14, $p < 0.01$, indicating a systematic response increase to the attended preferred SMS direction. The asymptote reflects the lowest response of the neuron, normally evoked by the anti preferred direction (moving in the opposite of the preferred direction, e.g. preferred direction = expansion, anti preferred direction = contraction). The mean Asymptote Index is 0.089 (Figure 8b). With an increase in firing rate across all directions of the attended tuning curve (Figure 9) we asked whether the neurons change their selectivity observed through a broadening or sharpening of the tuning curve. We did not observe a systematic change in tuning width (mean index = -0.006) $p = 0.8$ (Figure 8c) related to the spatially attentional location. The direction parameter (Figure 8d) plots the difference between the two preferred SMS directions attend inside the RF versus attend outside the RF. The axes represent the difference in degrees shifted positive or negative of the alignment to the preferred direction attended inside the RF. No systematic shift to either side was observed (3.8º, $p = 0.16$).

The distribution of preferred SMS directions along with the associated AI given in polar coordinates are depicted in Figure 10. The length of the individual direction vector marking the magnitude of the cells attentional modulation. More then half of the

recorded cells (approximately 60%) preferred stimuli composed of some combination of CW or CCW motion with expansion or pure expansion. This asymmetry in our MSTd directional preferences is consistent with previous reports (Heuer and Britten, 2004; Geesaman and Andersen 1996; Graziano, Andersen and Snowden, 1994; Saito, Yukie, Tanaka, Hikosaka, Fukada and Iwai, 1986; Tanaka and Saito 1989; Tanaka, Fukada and Saito, 1989). Further, we report that this anisotropy does not only typify the tuning preferences of MSTd but also the attentional modulation. The 73 cells responding preferentially to some form of expansion are significantly ($p < 0.05$) more modulated by spatial attention, nearly twice as much (37%, AI = 0.16) as cells responding to contracting SMS stimuli (21%, AI = 0.095).

Effects of attention based on MSTd cell types

Comparing the attentional modulation of glutamatergic pyramidal neurons and GABAergic interneurons (Markram, Toledo-Rodriguez, Wang, Gupta, Silberberg and Wu, 2004; Wonders and Anderson, 2006) did not lead to conclusive attention induced response differences. Responses of fast spiking interneurons with a mean trough to peak duration of 146.1 μs showed an spatial attentional enhancement of ~29% (Figure 11). The attentional response increase of broad spiking (mean trough to peak duration of 304.6 μs) pyramidal cells did not significantly differ ($p = 0.07$, Wilcoxon rank sum test) from the inhibitory interneurons (40% spatial attention response enhancement).

Discussion

Our data show that directing voluntarily attention to a spatial location, increases responses of MSTd cells coding that location. The observed response increase to an attended stimulus inside the RF over the identical unattended stimulus was 30%. We made use of optimized spiral motion stimuli (SMS) to evoke reliable responses to preferentially coded complex motion stimuli within area MSTd. The phenomenon described here bear on the important function of spatial attention as it modulates responses of neurons in higher level motion processing area MSTd twice as much as those in adjacent area MT an intermediate stage of motion processing. The bottom-up visual signal and the top-down attentional mechanism interact to amplify and transform the relevant information to create a saliency map accentuating regions of interest in the visual field (Treue, 2003). The observation that attentional modulation strength does depend on the preferentially coded direction of MSTd cells is not predicted by the

standard spotlight theory of attention. Nevertheless, our results provide some further insight to the recently reported discriminability of spiral motion in humans (Rodríguez-Sanchez, Tsotsos, Treue and Martinez-Trujillo, 2009).

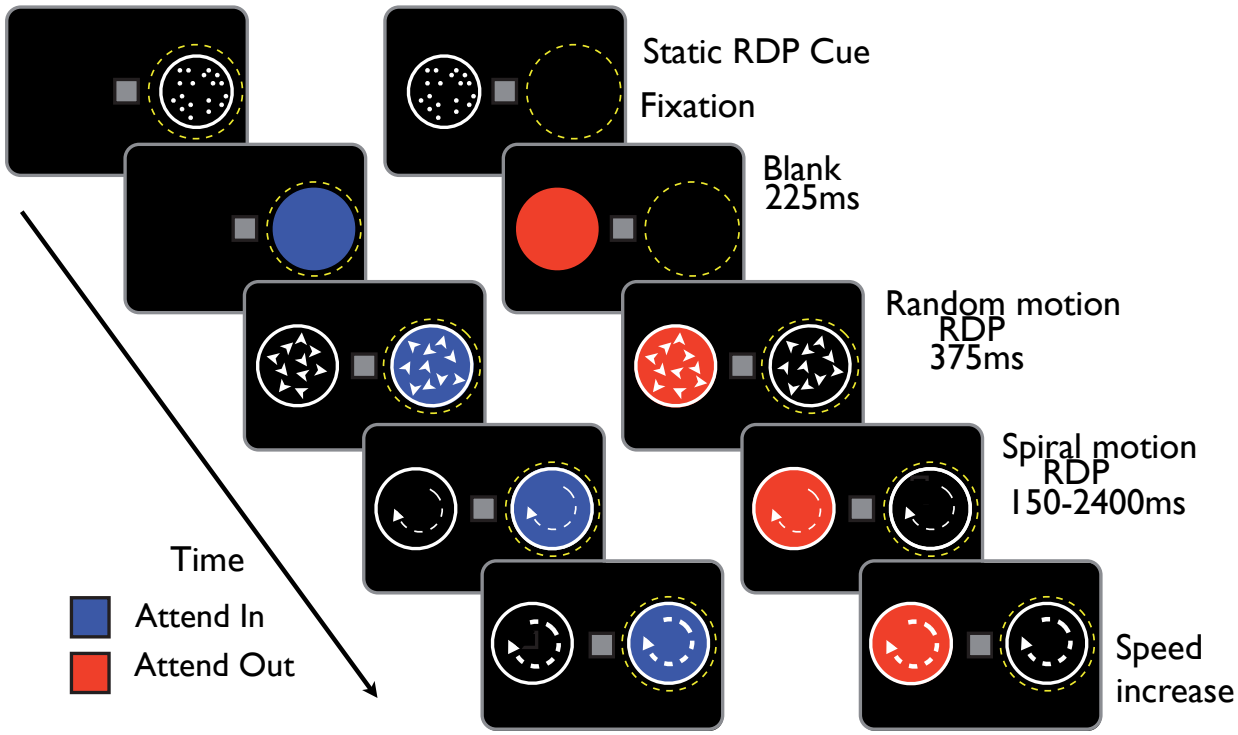


Figure 1: Schematic trial structure of the spatial attention task (attention inside RF / attention outside RF). At trial start both a small fixation spot, positioned in the center of the screen and a static RDP are presented simultaneously. The static RDP served as a cue to direct the animals attention to the target location and ceased to be visible once the animal fixated and initiated the trail by touching the lever. Following a 225 ms blank period, two RDPs were presented (displaying random motion) at equal eccentricity one inside the RF (indicated by the yellow dotted line), the other outside. The cued stimulus was the target the other was the “distractor”. The colored circles indicate the focus of attention. In the attention inside RF (blue) condition the animal was rewarded for responding (quick lever release) to a brief increase in speed of the coherent SMS motion pattern. Trials were aborted if the animal did not respond to the rapid (200 ms) target change, responded to the distractor speed increase or broke fixation. In the other condition (red) the animal was cued to attend the stimulus placed outside the RF

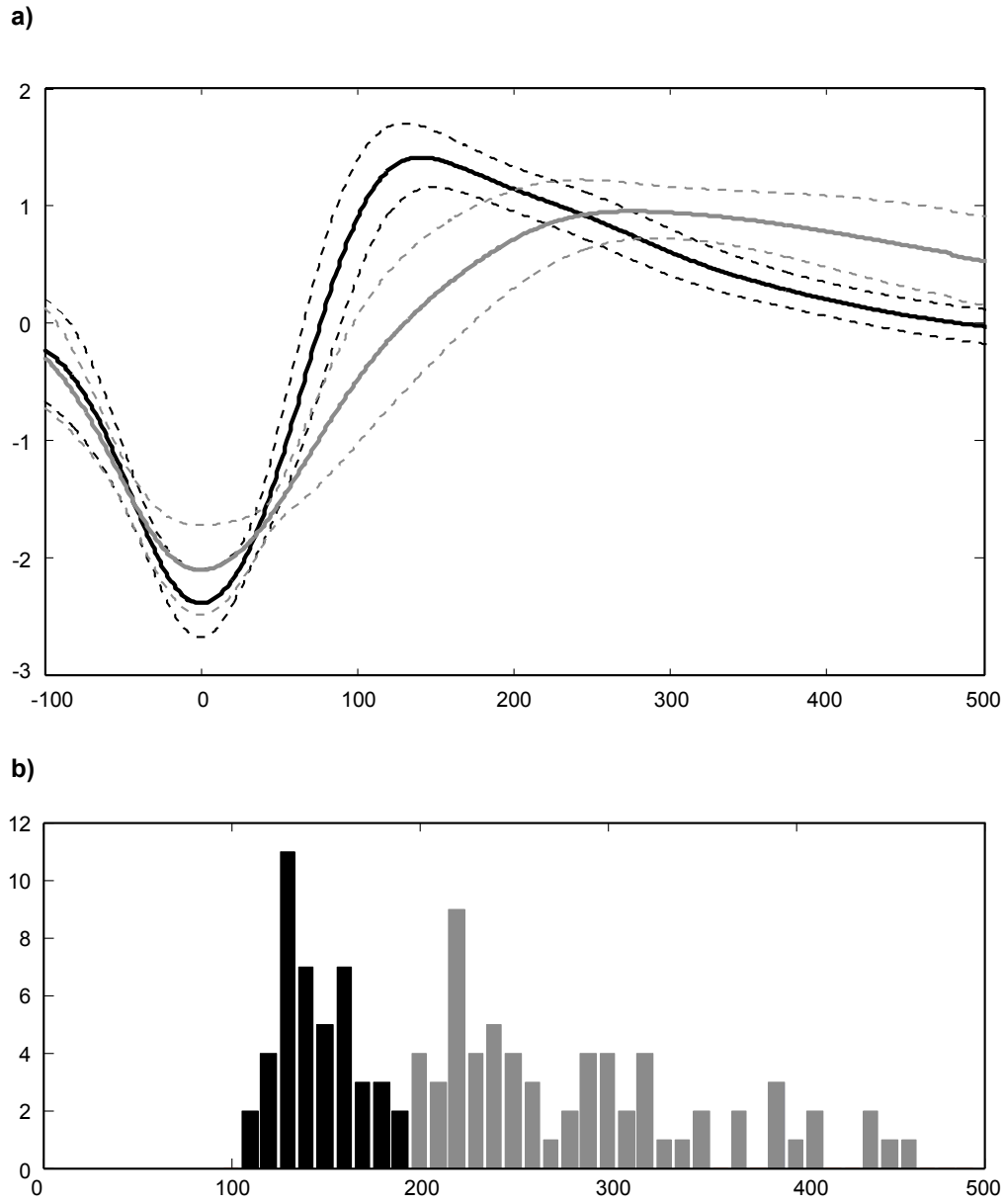


Figure 2: Classification of narrow and broad spiking neurons. (a) The mean waveforms of 121 neurons (±1 standard deviation); 44 narrow (black) and 77 broad (gray) waveforms of SMS responsive MSTd neurons. Waveform duration was measured as time from trough to peak in μs. (b) The bimodal waveform distribution.

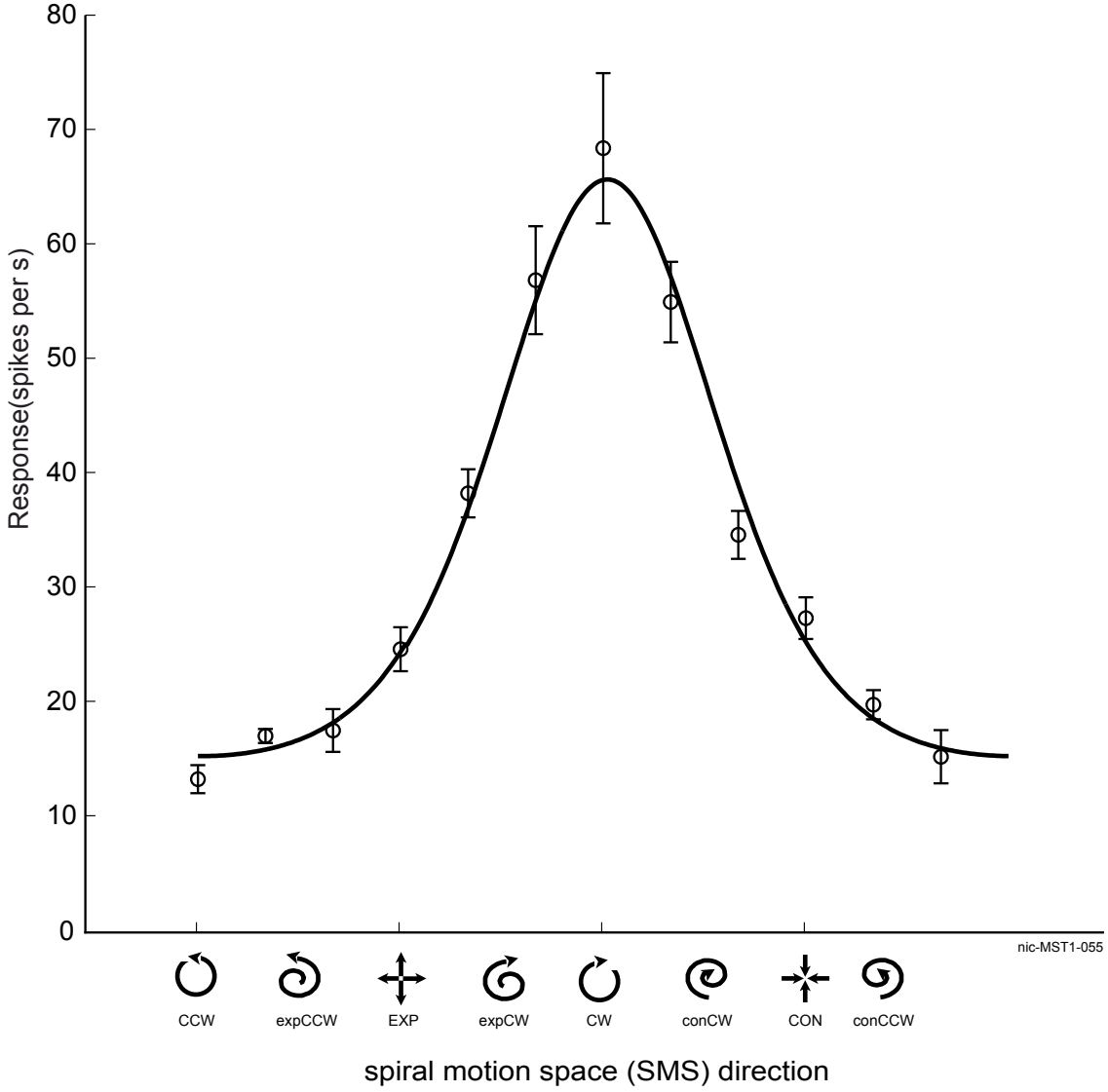
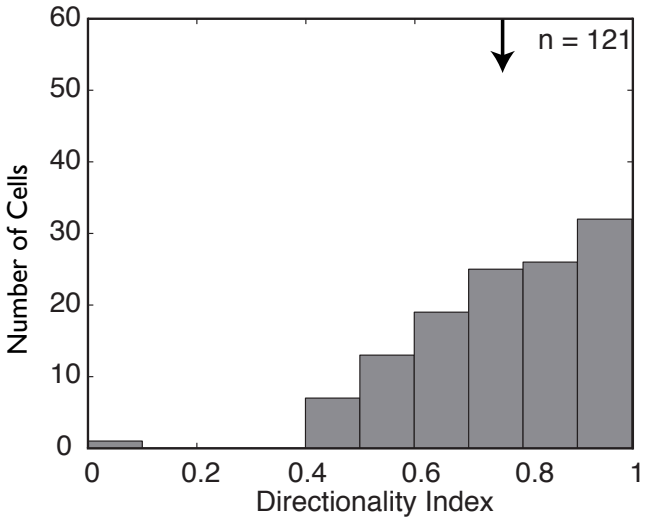


Figure 3: Tuning from one MSTd cell. Tuning curves, here for one example neuron were constructed fitting the responses of 12 SMS motion directions. The unattended activity of the cell was used to determine the tuning preferences of each of the 121 neurons.

a)



b)

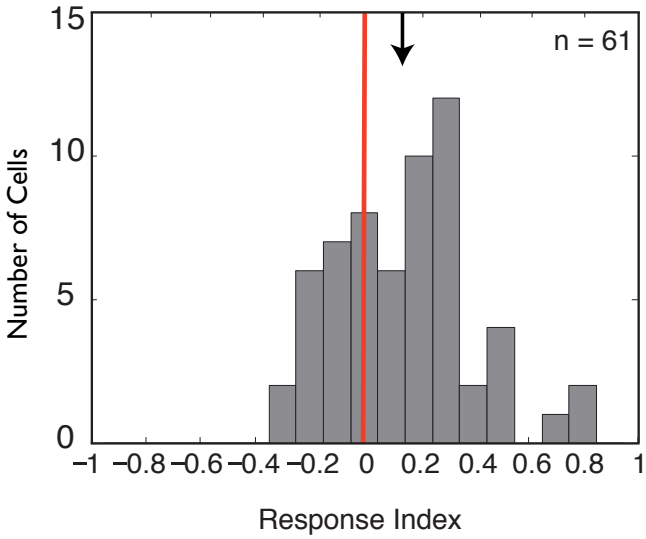


Figure 4: The tuning of MSTd cells to spiral motion stimuli is quite pronounced. This histogram (a) plots the directionality of the 121 cells we recorded from. The average index was 0.76. This Corresponds to a response to the preferred spiral space motion that is more than 4 times higher then the response to the anti preferred direction in spiral space. MST neurons are also tuned to linear motion with tuning to spiral (SMS) motion more pronounced. On average the responses are 33% higher (b) for spiral motion compared to linear motion.

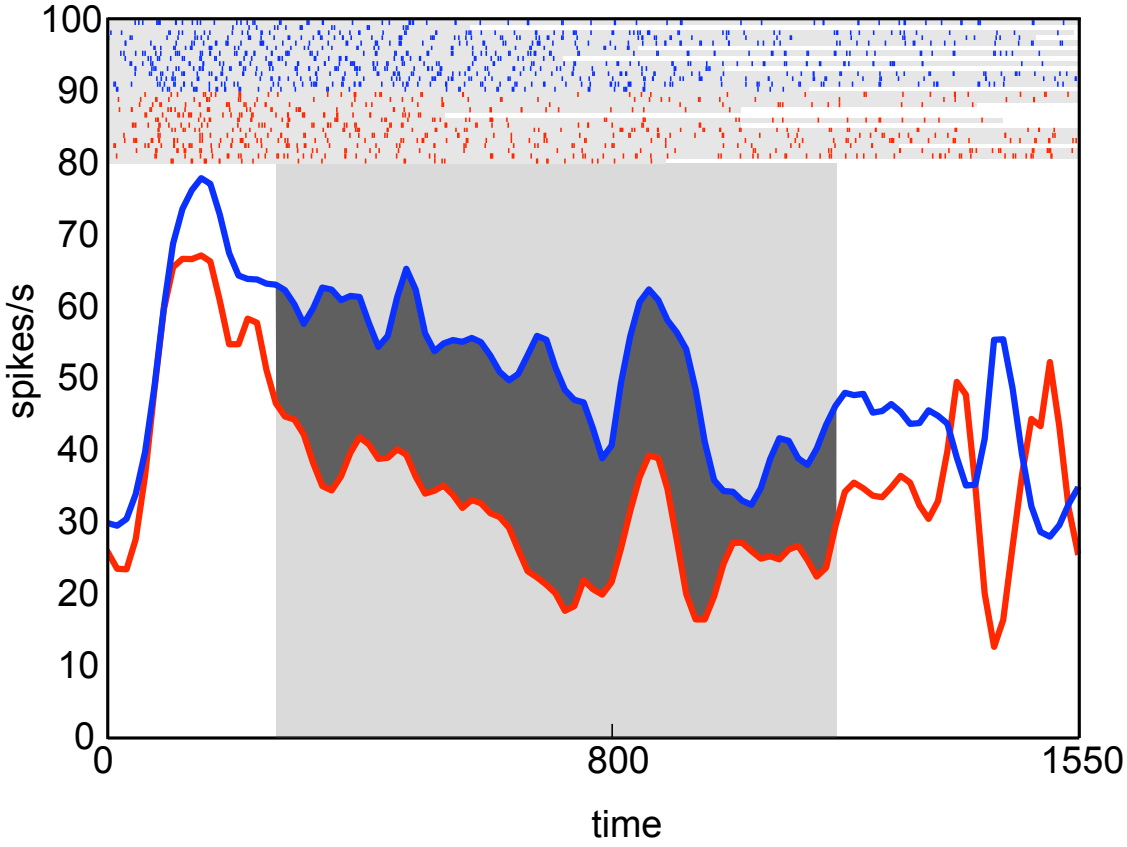


Figure 5: Effect of directing attention to the preferred spiral stimulus inside the RF (blue) or away from the identical stimulus (red). The activity of an example neuron. The modulation during the sustained period of attending a specific location is marked in gray.

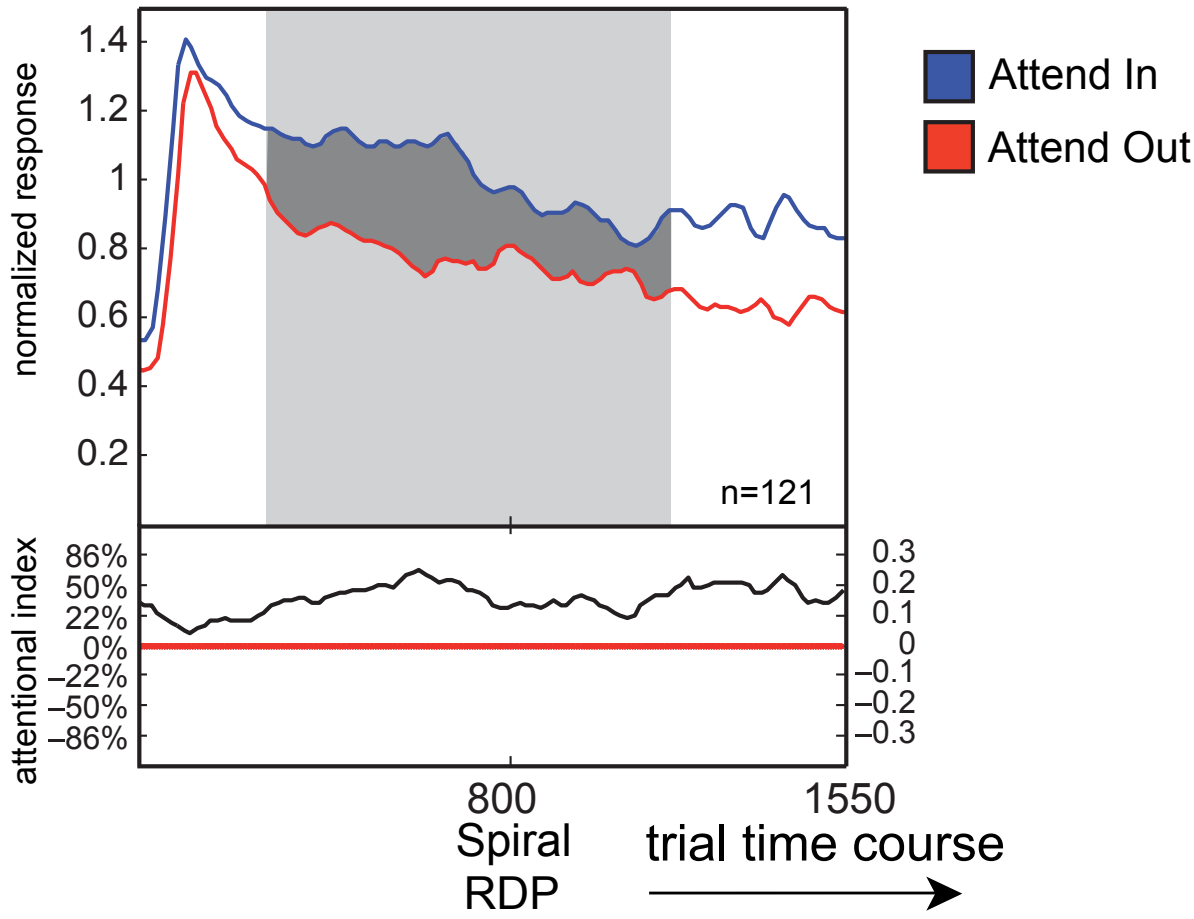


Figure 6: MSTd spatial attention time course. With the onset of coherent spiral motion a transient response onset can be observed. The gray box represents the 850 ms time window that we used for our analysis. Plot here the response to the preferred motion attended (blue) and unattended (red). Responses to the same sensory input when attention is inside the receptive field show a much higher response. On average this represented a 30% increase. The attentional modulation time course is plotted below showing the response modulation over time.

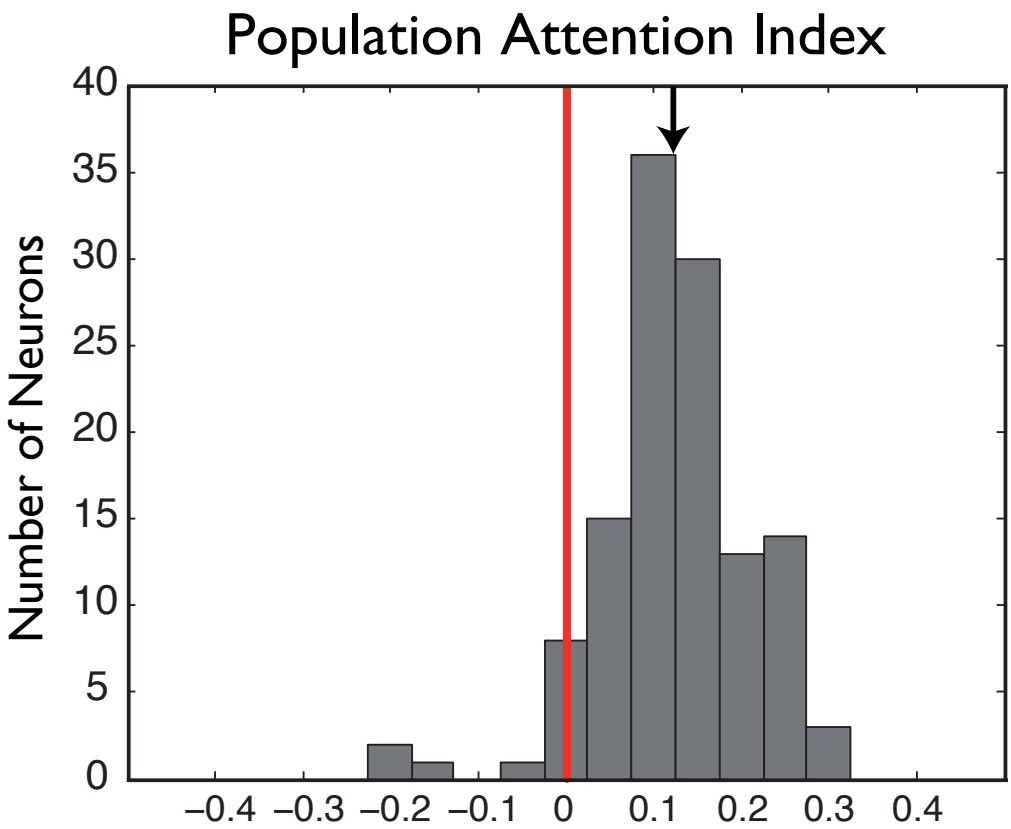


Figure 7: Distribution of the spatial attentional effect for the recorded population of 121 MSTd neurons. When directing attention to the preferred spiral motion inside the RF as apposed to attending away from the SMS motion stimulus inside the RF the AI is shifted to positive values. Indicating that attention modulates the processing of the MSTd spiral motion signal.

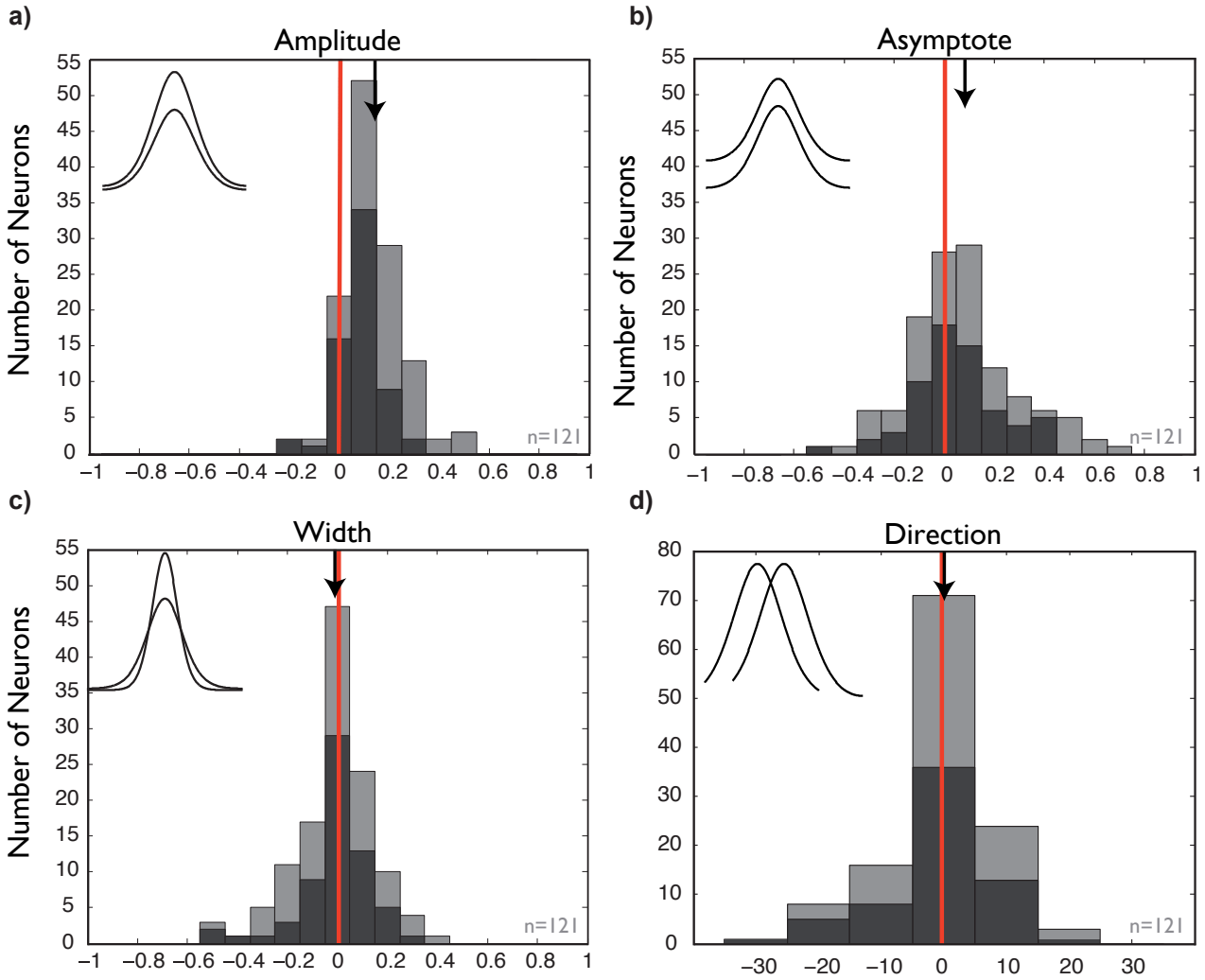


Figure 8: The changes in tuning curve parameters evoked by attention represented here using index values. (a) The amplitude, (b) asymptote, (c) width and (d) direction (displayed in degrees). Binning is based on the attentional index AI = ($response_{Attend\ IN}$ – $response_{Attend\ OUT}$)/($response_{Attend\ IN}$ + $response_{Attend\ OUT}$). Both top histograms ((a) amplitude and (b) asymptote) are shifted to positive values, indicating that spatial attention modulates processing spiral motion in MSTd. For width and direction (c and d) the corresponding histogram is centered on zero (red-line).

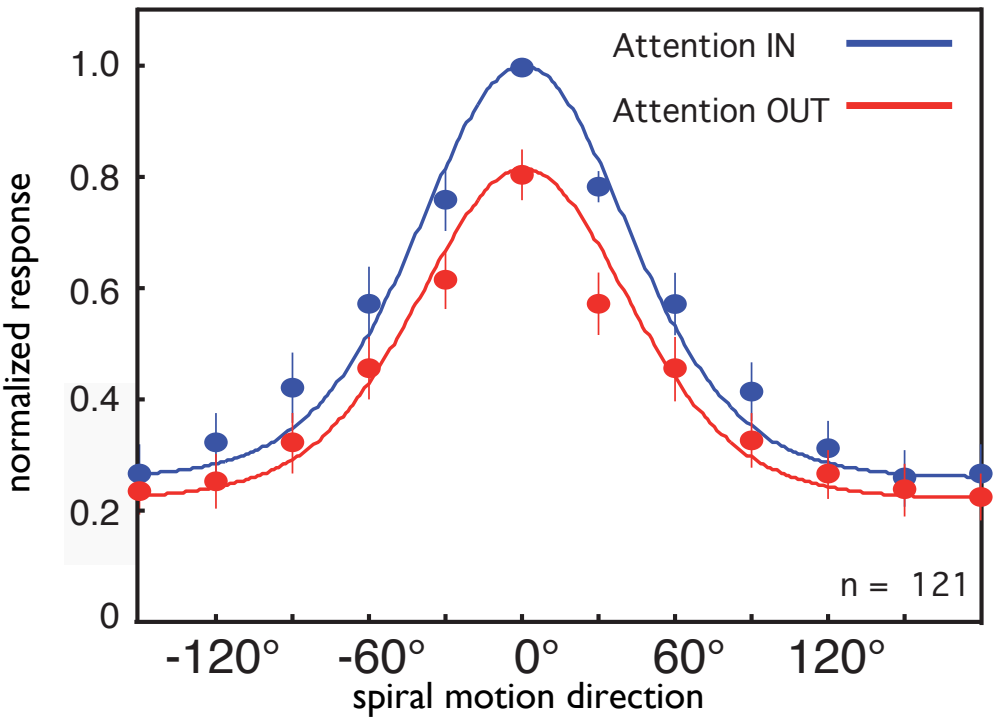


Figure 9: The population-tuning curves for 121 MSTd neurons that were tuned in both attended (blue circles) and unattended (red circles) task modes. The preferred orientation for each cell was aligned to the center taking the attended condition (red) as reference. Gaussians were fitted to the data and the averaged parameters (amplitude, asymptote and width) are displayed here together with the normalized data of 121 cells.

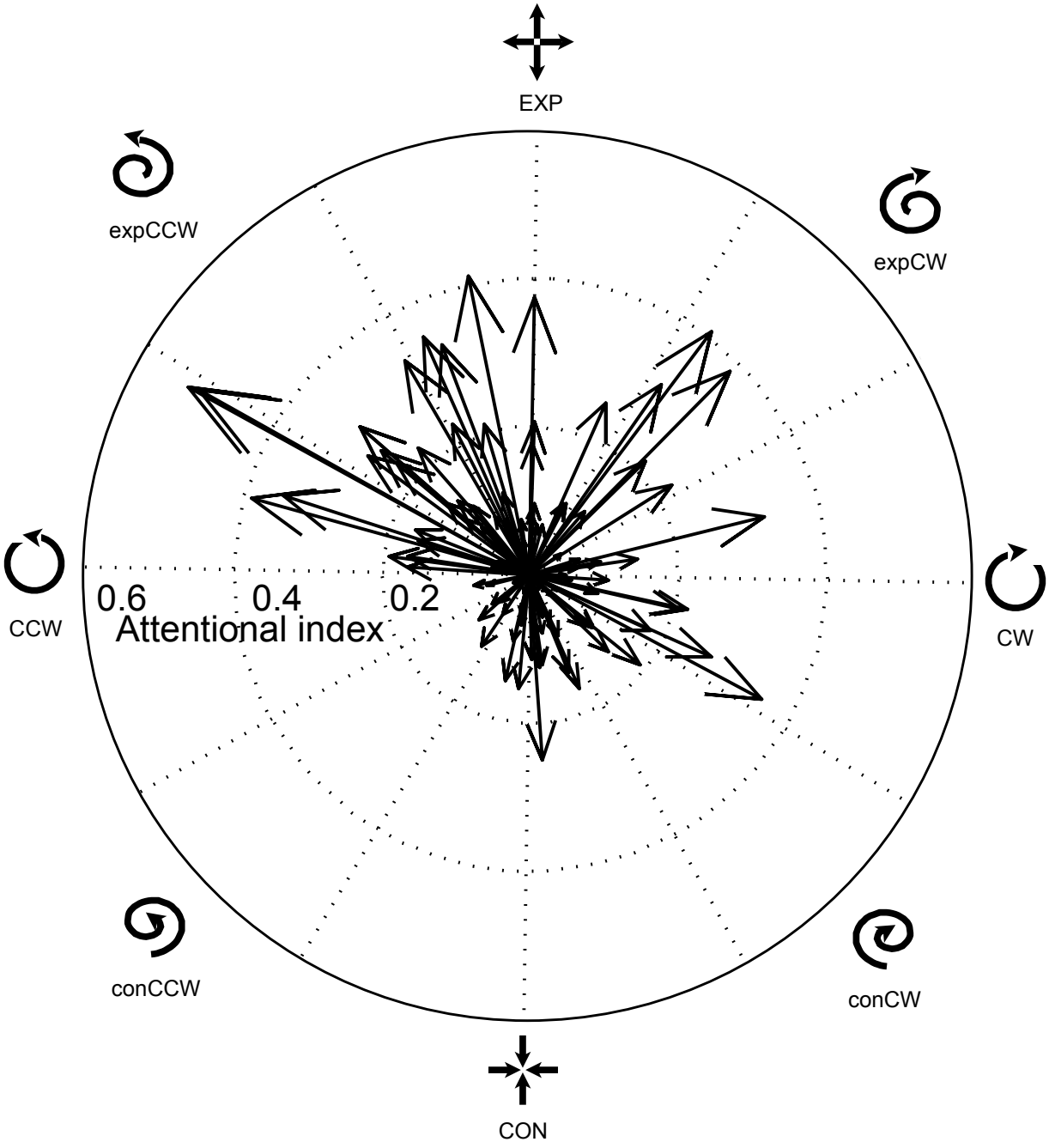


Figure 10: Preferred spiral directions for 121 cells. Each arrow represents one cell with the length of the arrow indication the attentional modulation of the individual preferred direction. The preferred direction of each cell was extracted from the Gaussian fit of the tuning curve.

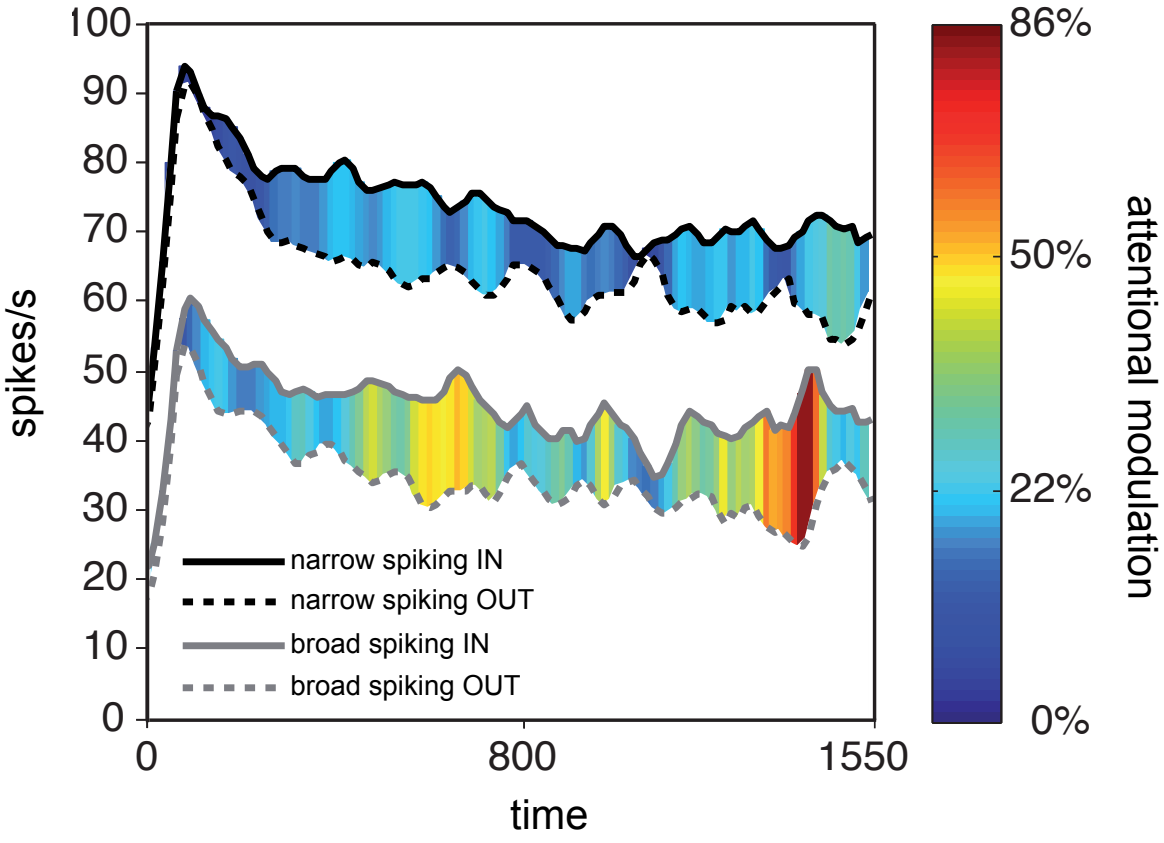


Figure 11: Average firing rates of broad spiking (gray) and narrow spiking (black) neurons. Both, for narrow and broad spiking neurons the attended condition (solid lines) shows significant modulation over the unattended (dashed lines) condition. The color index show the percent of positive modulation.

References

Cook EP, Maunsell JH. Dynamics of neuronal responses in macaque MT and VIP during motion detection. Nat Neurosci 2002; 5 (10): 985-94.

Duffy CJ, Wurtz RH. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J Neurophysiol 1991; 65 (6): 1329-45.

Duffy CJ, Wurtz RH. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. J Neurosci 1995; 15 (7 Pt 2): 5192-208.

Ferrera VP, Rudolph KK, Maunsell JH. Responses of neurons in the parietal and temporal visual pathways during a motion task. J Neurosci 1994; 14 (10): 6171-86.

Geesaman BJ, Andersen RA. The analysis of complex motion patterns by form/cue invariant MSTd neurons. J Neurosci 1996; 16 (15): 4716-32.

Graziano MS, Andersen RA, Snowden RJ. Tuning of MST neurons to spiral motions. J Neurosci 1994; 14 (1): 54-67.

Heuer HW, Britten KH. Optic flow signals in extrastriate area MST: comparison of perceptual and neuronal sensitivity. J Neurophysiol 2004; 91 (3): 1314-26.

Johnston K, DeSouza JF, Everling S. Monkey prefrontal cortical pyramidal and putative interneurons exhibit differential patterns of activity between prosaccade and antisaccade tasks. J Neurosci 2009; 29 (17): 5516-24.

Lagae L, Maes H, Raiguel S, Xiao DK, Orban GA. Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. J Neurophysiol 1994; 71 (5): 1597-626.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol 1997; 77 (1): 24-42.

McAdams CJ, Maunsell JH. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci 1999; 19 (1): 431-41.

McAlonan K, Cavanaugh J, Wurtz RH. Guarding the gateway to cortex with attention in visual thalamus. Nature 2008; 456 (7220): 391-4.

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. Interneurons of the neocortical inhibitory system. Nat Rev Neurosci 2004; 5 (10): 793-807.

Maunsell JH. The brain's visual world: representation of visual targets in cerebral cortex. Science 1995; 270 (5237): 764-9.

Maunsell JH, Cook EP. The role of attention in visual processing. Philos Trans R Soc Lond B Biol Sci 2002; 357 (1424): 1063-72.

Maunsell JH, van Essen DC. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. J Neurosci 1983; 3 (12): 2563-86.

Mitchell JF, Sundberg KA, Reynolds JH. Differential attention-dependent response modulation across cell classes in macaque visual area V4. Neuron 2007; 55 (1): 131-41.

Motter BC. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol 1993; 70 (3): 909-19.

Recanzone GH, Wurtz RH. Effects of attention on MT and MST neuronal activity during pursuit initiation. J Neurophysiol 2000; 83 (2): 777-90.

Rodríguez-Sanchez AJ, Tsotsos JK, Treue S, Martinez-Trujillo JC. Comparing neuronal and behavioral thresholds for spiral motion discrimination. Neuroreport 2009; 20 (18): 1619-24

Roy JP, Wurtz RH. The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. Nature 1990; 348 (6297): 160-2.

Saito HA, Yukie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J Neurosci 1986; 6 (1): 145-57.

Seidemann E, Newsome WT. Effect of spatial attention on the responses of area MT neurons. J Neurophysiol 1999; 81 (4): 1783-94.

Tanaka K, Fukada Y, Saito HA. Underlying mechanisms of the response specificity of expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. J Neurophysiol 1989; 62(3): 642-656.

Tanaka K, Saito HA. Analysis of motion of the visual field by direction, expansion/ contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. J Neurophysiol 1989; 62(3): 626-641.

Treue S. Neural correlates of attention in primate visual cortex. Trends Neurosci 2001; 24 (5): 295-300.

Treue S. Visual attention: the where, what, how and why of saliency. Curr Opin Neurobiol 2003; 13 (4): 428-32.

Treue S, Martínez Trujillo JC. Feature-based attention influences motion processing gain in macaque visual cortex. Nature 1999; 399 (6736): 575-9.

Treue S, Maunsell JH. Attentional modulation of visual motion processing in cortical areas MT and MST. Nature 1996; 382 (6591): 539-41.

Williford T, Maunsell JH. Effects of spatial attention on contrast response functions in macaque area V4. J Neurophysiol 2006; 96 (1): 40-54.

Wonders CP, Anderson SA. The origin and specification of cortical interneurons. Nat Rev Neurosci 2006; 7 (9): 687-96.

III.III.II Feature-based attentional modulation of the tuning of neurons in macaque area MSTd to spiral and linear motion patterns

Attention enhances the responses of neurons across the visual cortex, when directed to a behaviorally relevant location within the visual field (spatial attention) or a particular feature (feature-based attention). Previously research showed that allocating attention to a feature (e.g. direction of motion) creates a location-independent modulation of responses across the entire visual field, enhancing responses in the medial temporal area (MT) of macaque monkeys by roughly 13% when attention was directed to the preferred (feature) motion direction as compared to the non-preferred motion direction outside the receptive field (Treue and Martinez-Trujillo,1999). A feature similarity gain model was proposed, where the magnitude of attentional modulation of the response of a neuron reflects the similarity of the behaviorally relevant feature and the sensory selectivity of the neuron along this dimension. In the previous manuscript we demonstrated that allocating spatial attention to a particular location in visual space enhances the responses of neurons to spiral motion space (SMS: a continuum of motion stimuli in between expansion, clockwise rotation, contraction and counterclockwise rotation) stimuli within the dorsal region of the medial superior temporal area (MSTd) by 30 %; twice as strong as the spatial attentional modulation of responses of MT to linear motion.

Here we investigated feature-based attentional modulation in MSTd. We trained two rhesus monkeys to perform a feature-based attention task, attending to a cued direction of either SMS or linear motion. The monkeys had to maintain their gaze at a central fixation spot while a preferred SMS / linear motion stimuli was presented inside the receptive field and simultaneously in the opposite hemifield a cued target SMS / linear motion stimulus was moving either in the preferred or the non-preferred direction.
We compared the responses of 104 MSTd neurons well tuned to SMS and modulated by spatial attention in these two attentional conditions (feature-based attnetion: attend preferred versus attend anti-preferred). The observed response increased by 13 % with attention on the preferred feature versus attending the non-preferred feature. Our results show that feature-based attention multiplicatively modulates the response of MST neurons to SMS by the same factor as the responses of MT neurons to linear motion. Additionally we recorded the spatial - and feature-based attention tasks in a

sub-population of 43 neurons presenting linear motion stimuli. While we did observe a significant spatial attentional modulation of linear motion in the spatial attention task the MSTd neurons did not treat linear motion as a feature. The comparable increase for only spiral motion feature-based attentional modulation from area MT (linear motion) to MSTd (compared to the doubling of modulation observed for spatial attention for both linear and spiral motion) might reflect the appearance of SMS tuning only in MSTd while MST's spatial selectivity (and modulation by spatial attention) might be inherited from MT.

Feature-Based Attention in MSTd: not all stimuli giving rise to a response are treated as a feature

Daniel Kaping, Sonia Baloni, Stefan Treue
Cognitive Neuroscience Laboratory, German Primate Center, Goettingen, Germany
Bernstein Center for Computational Neuroscience, Goettingen, Germany

Our ability to respond fast (Posner, Snyder and Davidson,1980) and accurately (discrimination sensitivity, Lu and Dosher, 1998; contrast sensitivity, Cameron, Tai and Carrasco, 2002; distractor interference, Shiu and Pashler, 1995; acuity, Yeshurun and Carrasco, 1998) to a stimulus can be greatly enhanced by directing attention in advance to the target position. Attention operating on a spatial location, describes the location-specific focus, implied by the term “spotlight of attention” (Eriksen and Hoffman, 1973; Eriksen and Eriksen, 1974; Posner, Snyder and Davidson,1980; Hurlbert and Poggio, 1985), of highly efficient processing. This should operate and enhance all properties of the incoming sensory signal (bottom-up) within the “spotlight” while relegating everything beyond the focus of spatial attention into the shadow of conscious perception. Several single unit studies have reported substantial spatial attentional modulation of neural activity in specific cortical areas, along both the ventral (Moran and Desimone, 1985; Motter, 1993; Desimone and Duncan, 1995; Connor, Preddie, Gallant and Van Essen, 1997; McAdams and Maunsell, 1999) and dorsal (Treue and Maunsell, 1996; Cook and Maunsell, 2002) visual pathways, of stimuli perceptually analyzed within the particular region. Neurons respond stronger to the physically identical stimulus, placed within the receptive field, when attention is directed into the receptive field as compared to attending elsewhere in the visual field. Complementary spatial attention induced increases in brain activity have been reported by human neuroimaging studies (Corbetta, Miezin, Shulman and Petersen, 1993; Tootell, Hadjikhani, Hall, Marrett, Vanduffel, Vaughan and Dale, 1998; Brefczynski and DeYoe,1999; Gandhi, Heeger and Boynton, 1999). The strength of the (top-down) spatial attentional effect onto the sensory signal propagating through the visual processing hierarchy increases (Luck, Chelazzi, Hillyard and Desimone, 1997; Cook and Maunsell, 2002; Maunsell, 2004)

along with the complexity of the preferentially coded stimulus attributes. While spatial attention strongly impacts visual processing and therefore perceptual representation, prior knowledge about the location of a behavioral relevant stimulus does not always exist. A mechanism (feature-based attention) employing a particular attribute (feature) of the behavioral relevant stimulus, such as color or direction of motion, dissociates attention from a specific location, enhancing processing of the attended feature in the functionally specialized visual areas across the entire visual field. Feature-based attention is exceptionally useful in visual search paradigms, in which an observer is instructed to find a predefined target item embedded in a display of distractors. All potential targets should selectively be enhanced to create a feature-based saliency map (Treue, 2003b) on which the visual focus can then be oriented. Several studies have examined the effects of feature-based attention. Neuronal responses of orientation selective neurons in area V4 increase their response by 11% when the animal attended to the orientation versus the color of a stimulus outside the RF of the recorded neuron (McAdams and Maunsell, 2000). Comparable results of feature-directed attention modulated responses of neurons in the middle temporal area (MT) have been reported by Treue and Martínez-Trujillo (1999). They found responses to be 13% stronger for the unattended preferred motion stimulus inside the RF when the animal attended to a stimulus placed outside the RF moving in the identical, preferred direction of the neuron versus a stimulus moving in the anti-preferred, 180º opposite direction. These findings have lead to the feature-similarity gain hypothesis of attention (Treue and Martínez-Trujillo, 1999) where the strength of attentional modulation depends upon the similarity of the attended feature and the preference of the neuron. Since both, feature-based and spatial attention coexist and seem to act independently (Treue and Martínez-Trujillo, 1999), but with similar mechanisms, it remains unknown if and to what extent it is passed-on along the visual processing hierarchy?

To answer this question we measured the spatial and feature-based attention induced response modulation of cells in the dorsal region of the medial superior temporal area (MSTd). Neurons in MSTd are preeminent to investigate not only the propagating motion signals, as MSTd receives direct input from area MT (Ungerleider and Desimone, 1986) but additionally the top-down attentional influences at the next processing level (ascending from MT) of the bottom-up sensory information. Besides responding to linear motion in similar fashion as area MT, MSTd neurons preferentially code more complex spiral motion patterns (Saito, Yukie, Tanaka, Hikosaka, Fukada and

Iwai, 1986; Tanaka, Hikosaka, Saito, Yukie, Fukada and Iwai, 1986; Duffy and Wurtz, 1991; Orban, Lagae, Verri, Raiguel, Xiao, Maes and Torre, 1992; Graziano, Andersen and Snowden, 1994; Geesaman, Andersen, 1996). Thus we can make comparisons to the well reported attention related response changes of linear motion in area MT (feature-based and spatial attention) measuring MSTd responses with and without attention in the receptive field (spatial attention, linear and spiral motion); in addition to directing attention to a particular motion feature (linear and spiral motion) away from the receptive field. If MSTd neurons indeed inherit spatial information from prejacent area MT responses across both measured dimensions, spiral and linear motion should be enhanced in a similar fashion. As spatial attention operates in a multiplicative fashion it should enhance any stimulus triggered response by the same multiplicative factor. However some important issues remain. How does the linear motion selectivity of the visual system (e.g. area MT) contribute to MSTd linear and spiral motion responses? Does feature-based attention modulate MSTd responses? How comparable are these, as linear motion is already processed in area MT and MSTd is the first known place within the primary visual cortex to encode more complex spiral motion patterns.

In a previous experiment we found that attending to a stimulus positioned in the RF caused a consistent 30% response increase in area MSTd. Here we observe a multiplicative response enhancement of 30% that is invariable regardless of presenting spiral or linear motion during the spatial attention task; which is twice as strong as that reported for area MT (Treue and Maunsell, 1996; Katzner, Busse and Treue, 2009). Expecting the linear motion signal to be passed on from area MT we were surprised to find no feature-based attentional modulation to linear motion in MSTd. To assure that we did not distort our linear motion feature-based attention MSTd results through poor stimulus placement we presented linear motion patches (of varying size) at different locations within the RF. MSTd cells maintained their linear motion selectivity at different positions in the RF allowing us to rule out position dependent effects of linear motion feature-based attention. Contrary to linear motion, spiral motion indeed showed that directing attention to a spiral motion feature did modulate the responses of neurons throughout the visual field.

Methods

All experimental procedures were approved by the local animal research committee and complied with relevant laws and institutional guidelines. Headpost and recording chamber were implanted using standard surgical techniques (see page 66 (Spatial attention modulates MSTd responses to optimized spiral motion space stimuli; Kaping, Baloni, Zhang and Treue).

Monkeys were seated in a custom-made primate chair with their head restrained at a distance of 57 cm from a computer monitor (resolution 40 pixels per degree of visual angle, refresh rate 76 Hz). The eye position was monitored (sampling rate of 230 Hz) with an infrared video-based eye tracker (ET49, Thomas Recording, Giessen, Germany). Visual stimuli consisted of random dot patterns (RDPs) moving within a circular stationary virtual aperture. The size of the RDP, the speed of the dots, and the direction of motion were selected to match the preferences of the MSTd neuron under study.

Stimuli

We used random dot patterns (RDPs) of small bright dots (density: 8 dots per degree^2, luminance 75 cd/m^2) plotted within a stationary circular aperture on a gray background of 35 cd/m^2. Stimuli were spiral motion space (SMS) patterns or fully coherent linear motion patterns. SMS considers expansion, clockwise rotation, contraction and counterclockwise rotation as neighboring stimuli with a continuum of stimuli in between these cardinal directions.

Behavioral Task

Two rhesus monkeys (Macaca mulatta) were trained to attend either to the spiral motion space (SMS task) or to the unidirectional linear motion (LM task) signal of moving random dot pattern (RDP) (Figure 1). The animal initiated a trial by maintaing fixation within a window of 1.5º radius, centered on a 0.2º fixation square. The color of the fixation point indicated whether a fixation (white - sensory condition) or feature-based / spatial attention task (black) was upcoming. In the fixation task, two moving RDPs were presented at equal eccentricity in opposite visual hemifields, one of them inside the receptive field (RF) of the MSTd neuron under study. The animal received a liquid reward for releasing the lever within a response time window of 100–500 ms following a brief luminance change on the central fixation point while not responding to any

changes of the RDPs. To ensure that the monkeys were correctly attending to the luminance change either of the two RDPs, randomized in time and order, could increase their speed. In the feature-based / spatial attention task a stationary RDP cue appeared at the target location for 80 ms signaling the position of the upcoming relevant stimulus. The monkey received a liquid reward for releasing the lever within a response time window (identical as for fixation trials) following a brief speed increase of the cued RDP while ignoring all changes within the uncued location ('wrong location' distractor). Responses to distractor events at the wrong location allowed to control if the monkeys attended to the target location. Trials were ended immediately following any response, obviously no reward was given for wrong responses and fixation errors.

Spatial Attention Task

Two SMS stimuli moving in the preferred direction of the neuron were simultaneously presented: (i) one inside the RF, (ii) the second in the opposite hemifield. The presentation ensures that feature-based attention is equated between the pair of physically identical stimuli. The animal had to reported a speed change within the target stimulus (either inside the RF or outside the RF) while ignoring all changes within the other, distractor stimulus (Figure 1).

Feature-Based Attention Task

We have trained two rhesus monkeys to perform a feature-based attention task, attending to a cued direction of SMS in either preferred direction or the non-preferred direction. The monkeys had to maintain their gaze at the central fixation spot while a preferred SMS stimuli was presented inside the receptive field and simultaneously in the opposite hemifield a cued target SMS stimulus varied between the preferred and the non-preferred direction (Figure 1).

Linear Motion Position Invariance

While the animal engaged in a luminance change detection task on the fixation point we tested individual MSTd cell linear motion selectivity regardless of stimulus position and size in the cell's receptive field. A set of five stimuli was used to test linear motion selectivity across the receptive field. While one unattended large field linear motion stimulus, matching size and location of the attention task stimuli was used as a reference, four smaller patches of linear motion tested the direction selectivity across the receptive field. The position lay along the horizontal and vertical axis with respect to

the center of the receptive field center. Each motion patch covered 1/4 of the diameter of the large motion stimulus used as reference (positioned identical to stimuli used throughout the attentional tasks) to determine the position-dependent linear motion responses. The individual motion patches as well as the large reference stimulus moved in a random unidirectional motion sweep, changing direction every 300 ms. While the transient onsets of constantly changing motion stimuli was disregarded, the sustained response of the last 200 ms of each individual presented motion direction was analyzed. Responses of the neurons were sampled for 12 directions in steps of 30 degree and fitted with a Gaussian function to obtain position-specific tuning profiles.

Neuronal Recording and Data Collection

Single-unit activity was recorded from area MSTd using a three-channel multi-electrode recording system (Mini-matrix, Thomas Recording, Plexon data acquisition system). For all recording sessions, three electrodes were simultaneously advanced to isolate individual MTSd. Cells were characterized as MSTd neurons based on spiral motion directional tuning, receptive size, and position in the cortex. The locations and sizes of individual RFs were determined by manually moving a moving RDP. Spiral motion space directional tuning was determined by presenting a single RDP inside the RF, moving in 12 different SMS directions; while linear motion was determined by presenting 12 different unidirectional linear motion directions. Preferred direction was estimated by fitting a Gaussian to the individual neuronal responses, plotted as a function of stimulus direction. Each preferred motion stimulus (SMS and LM) was presented in 8 different speeds (0.5–64 deg/s) to determine speed tuning, while monkeys were maintaining fixation.

Data Analysis

Data were analyzed offline with custom scripts using MATLAB (The Math Works Inc., Natick, MA). For the analysis of neuronal data only correctly performed, completed trials were included. Spike density functions (SDF) were obtained by convolving each spike with a Gaussian function ($\sigma = 20$). Response rates were determined by averaging the frequency of action potentials over 800 ms of stimulus presentation starting 260 ms after coherent motion onset.

Population responses were computed averaging across normalized responses. For each recorded neuron the differences in response between the attended and

unattended condition were computed (AI (attentional modulation index : [(attention-in - attention-out)/(attention-in + attention-out)])). The AI is a conservative measure of the average attentional effect, reducing the effects of outliers. In the spatial attention task only trials in which the RDPs (inside and outside the RF) moved in the preferred direction of the neuron were included in this analysis. The statistical significance of the AI was evaluated using a one-sample t test. Spatial attention of linear motion was compared to the SMS condition using a Wilcoxon rank sum test.

The effects of feature-based attention were analyzed during the sustained responses in a time window of 260 ms - 1060 ms after the onset of coherent SMS. For each recorded neuron the changes in response between attending the preferred motion direction and attending the anti-preferred motion direction (both presented outside the RF) were computed (AI (attentional modulation index : [(attend preferred - attend anti-preferred)/ (attend preferred + attend anti-preferred)])). The statistical significance of the AI was evaluated using a one-sample t test. Feature-based attention of linear motion was compared to the SMS condition using a Wilcoxon rank sum test.

The position invariance was tested across five positions. One large field probe stimulus used as a reference stimulus matched size and position of the stimuli used troughout the attention tasks. The selectivity of all probe stimuli was measured by calculating the individual directional selectivity (directional selectivity = 1 - (response to anti-preferred linear motion direction / response to preferred linear motion direction)) at each of the five positions. Only cells that showed 3 times higher responses for the preferred direction over the anti-preferred directions across all five positions (Figure 6a, 24 of 30 cells) were used to measure the difference in mean directions. The absolute mean directional shift across 96 small probe stimuli in respect to the individual large motion stimuli was 12.1º. Measuring the directional shift across all five positions of each individual cell did not reveal a significant difference ($p = 0.65$).

Results

We examined the spatial and feature-based attentional effects on SMS stimuli of 104 isolated neurons in area MSTd in two monkeys. Most of the neurons showed a significant spatial attention effect. This can be observed in the sustained response period to coherent spiral motion moving in the preferred direction of the individual neuron (Figure 2a). We calculated an attention index, (attended response - unattended

response) over the sum of the two responses. The population of neurons showed consistently stronger responses when the preferred SMS stimulus in the receptive field was attended versus when the identical stimulus was presented as a distractor. Notably, the mean of the attentional index (Figure 2c) is shifted away from zero (0.13) corresponding to a 30% increase in activity. The significance was statistically evaluated using a one-sample t test ($p < 0.01$). For a subset of 43 neurons both SMS and linear motion spatial attention effects were measured in separate trials, presented in randomized order. The 43 neurons were well representative of the entire population of 104 neurons recorded in the SMS spatial attention task; as their attentional index did not significantly differ from the population (mean 0.128, Wilcoxon rank sum test, $p = 0.86$). The sub-population that we recorded showed not only the positive attention effect for SMS stimuli moving in the preferred SMS direction of the MSTd neuron but for also linear motion stimuli (Figure 2b) moving in the preferred linear direction. The mean attention index during the linear motion MSTd spatial attention task was 0.14, corresponding to a 33% ($p < 0.01$) increase in activity (Figure 2d). This suggests that spatial attentional modulation operates uniform along all stimulus dimensions as the two spatial attentional effects (linear and SMS) are not significantly different ($p = 0.96$). The magnitude of attentional gain for the SMS as well as the linear motion is consistent with the previously reported spatial attention induced response enhancements of MSTd (see page 71 (Spatial attention modulates MSTd responses to optimized spiral motion space stimuli; Kaping, Baloni, Zhang and Treue)).

During the feature-based attention task the animal was required to attend to the motion direction (feature) of the stimulus placed outside the RF either: (i) moving in the preferred direction of the neuron, or (ii) moving in the anti-preferred direction. To create a good response level from the neuron, the unattended stimulus, placed inside the receptive field was always moving in the preferred direction. This allowed us to observe the modulation within the network based upon neuronal selectivity to the attended motion feature. We tested the entire set of 104 neurons on SMS stimuli and the subset of 43 neurons (previously analyzed for spiral and linear motion in the spatial attention task) again presenting not only spiral motion but additionally linear motion in the feature-based paradigm. The responses of two single units are plotted in Figure 3. The visual stimulation was different as Figure 3a shows the response to SMS stimuli and Figure 3b that of linear motion; for both, attention was directed to the stimulus outside the neuron's receptive field switching between the preferred and the anti-preferred direction. Though

attending to spiral motion as a feature modulates the unit in Figure 3a,c surprisingly the unit in Figure 3b,d showed no obvious effect during the linear motion feature attention task. The population response time courses of coherent moving SMS stimuli (Figure 4a) versus linear motion stimuli (Figure 4b) behaved in the same dissimilar fashion as observed in the two single units. For the SMS feature-based attention task units responded more strongly when the animal was attending to a stimulus moving in the preferred SMS direction outside the receptive field. Although the response change (of roughly 11 %) is significantly smaller compared to the enhancement observed during a spatial attention task in MSTd, it is comparable with feature-based attentional enhancements previously reported in other visual areas (MT, Treue and Martínez-Trujillo, 1999; V4, McAdams and Maunsell, 2000). No significant effect was observed when the animal was directing its attention to the linear motion direction (Figure 4b). Calculating the attentional index for each neuron enabled us not only to statistically evaluate feature-based attention for SMS and linear motion separately but to qualitatively compare both conditions in neurons recorded with both tasks. The mean attention index during the spiral motion feature-based attention task (Figure 4c) was 0.05, corresponding to a ~11% increase in activity. While neurons within the subset significantly increased their response while the animal attended the preferred SMS direction attending to the preferred linear motion did not significantly ($p = 0.85$) modulate the population response (Figure 4d). Comparing feature-based attentional response modulations between SMS and linear motion trials within the same population of cells shows a significant difference (Wilcoxon rank sum test, $p < 0.05$). Although we selected cells randomly to be recorded with both linear and SMS motion, we did not consistently find MSTd cells that showed significant linear motion spatial and feature based attention enhancements. Remarkably most MSTd cells enhancing their response during the spiral motion spatial attention task also showed significant feature-based attentional enhancements.

Responses of linear motion tuning curves were tested with a motion sweep stimulus, changing unidirectional directional motion in a random order every 300 ms. One stimulus of equal size and position as the stimuli used in the attention condition was used as a reference to compare the obtained linear motion tuning curve at five different positions within the RF. Figure 5 shows the result for a MSTd cell preferring clockwise contracting SMS motion in addition to right downward linear motion. The cell maintained its preferred direction regardless of stimulus size and position as smaller probe stimuli

covered only one fourth of the large reference stimulus in size and were placed at four different positions within the RF. While some cells showed some form of directional displacement the overall selectivity did not change (Figure 6b). The mean of the position invariance measured cell by cell showed no significant difference between the large probe stimulus and the four smaller test probes (Wilcoxon rank sum test, $p = 0.65$, as the preferred direction of each of the four stimuli was aligned to the reference stimulus choosing).

Discussion

These experiments contribute to understanding how spatial and feature-based attention propagate in visual processing. They show that spatial information is inherited, enhancing the visual signal at the attended location. This finding is consistent with previous reports of spatial attentional (Treue and Maunsell, 1996). Further, we demonstrate that the same MSTd neuron selectively responding to changing types of visual input (unidirectional linear and complex spiral motion) strengthens its response by one multiplicative factor when the stimulus inside the neurons RF is attended. Both responses for spiral motion as well as linear motion were increased by roughly 30% when the animal attended the given motion stimulus. While these response enhancements are stimulus independent (linear and spiral motion) the linear motion spatial attentional modulation of MSTd allow for a direct comparison to spatial attentional modulation of area MT. As spatial attention effects have been shown to increase from “lower” levels to “higher” levels of visual processing (Luck, Chelazzi, Hillyard and Desimone, 1997; Cook and Maunsell, 2002; Maunsell, 2004) we are able to show that the modulation in area MSTd is doubled from prejacent area MT for stimuli in the same dimension (linear motion, adjusted to fit the specific MSTd neuron at hand, in stimulus size and preferred speed). The approximate doubling in response enhancement while attending a particular location in space from area MT (~ 15%, Treue and Maunsell, 1996) onto MSTd (30%) might be pure coincidental but an alternative explanation is that the bottom-up signal is modulated by comparatively the same factor along each processing step. At the earliest stages of visual processing mainly stimulus driven representation of the visual information (sensory signal) is modulated by a spatial attention factor (LGN, McAlonan, Cavanaugh and Wurtz,2008; V1, McAdams and Reid, 2005) resulting in a fraction of the enhancement observed later within the visual processing hierarchy. As this signal at the location of interest is passed along to the next processing levels not only the complexity of the preferential coded stimulus aspects

increase but also the top-down spatial attentional influence creates a more behavioral goal oriented representation of the visual information.

However, the increasing complexity of coded stimulus attributes raise the question as to what extend does the feature selectivity from one processing step to the next contribute and lives on. In our feature-based attention experiments, the process of attending to different types of visual motion input (spiral or linear) affected the signals of neurons throughout the visual field only for the motion signal to be coded / processed for the first time within the visual motion processing chain. MSTd is the first known processing step of complex spiral motion and attending to a particular spiral motion direction selectively increases the responses of all neurons sharing the same preferred spiral motion. As these MSTd spiral motion feature-based attention response enhancements of 11% are in magnitude comparable to that of linear motion feature-based attention in area MT (arguably the first specialized motion processing area) it might very well represent the expected absolute magnitude of attending to a discrete feature. Interestingly, while MSTd displayed highly selective responses to linear motion; attending to a given linear motion direction did not result in significant feature-based attentional enhancements across the recorded neuronal population. Along with our results of linear motion position invariance (that are complementary to the MSTd spiral motion position invariance reported by Graziano, Andersen and Snowden, 1994) it suggests that MSTd might not simply make use of the linear motion signal passed along from area MT but that it might have some intrinsic process generating linear motion selectivity.

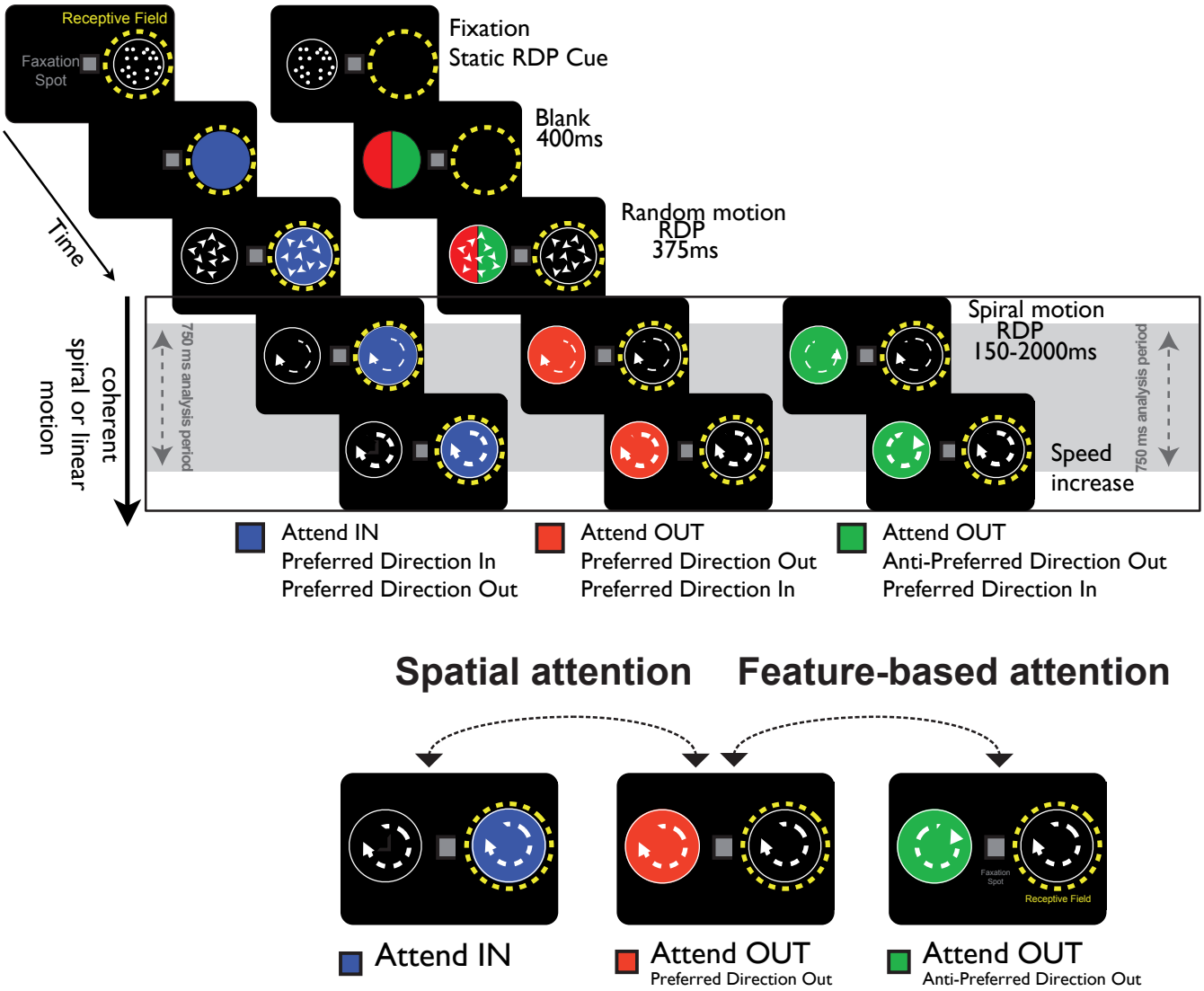


Figure 1: Schematic trial structure for the three experimental conditions (spatial attention / feature-based attention). After the animal fixated a small fixation spot, positioned in the center of the screen a static RDP flashed for 80 ms. The RDP served as a cue to direct the animals attention to the target location. Following a 400 ms blank period, two RDPs were presented (displaying random motion) at equal eccentricity one inside the RF (indicated by the yellow dotted line), the other outside. The cued stimulus was the target the other was the “distractor”. The colored circles indicate the focus of attention and the associated task. In the attention in (blue) condition the animal was rewarded for responding to a brief increase in speed of the coherent motion pattern (either spiral motion or linear motion). Trials were aborted if the animal did not respond to the target change, responded to the distractor speed increase or broke fixation. In the feature-based attention condition the animal was cued to attend the stimulus placed outside the RF (red and green). For both attend preferred direction of motion (red; either spiral or linear) and attend anti-preferred direction of motion (green; either spiral or linear) the stimulus placed inside the RF was moving in the preferred direction of the cell.

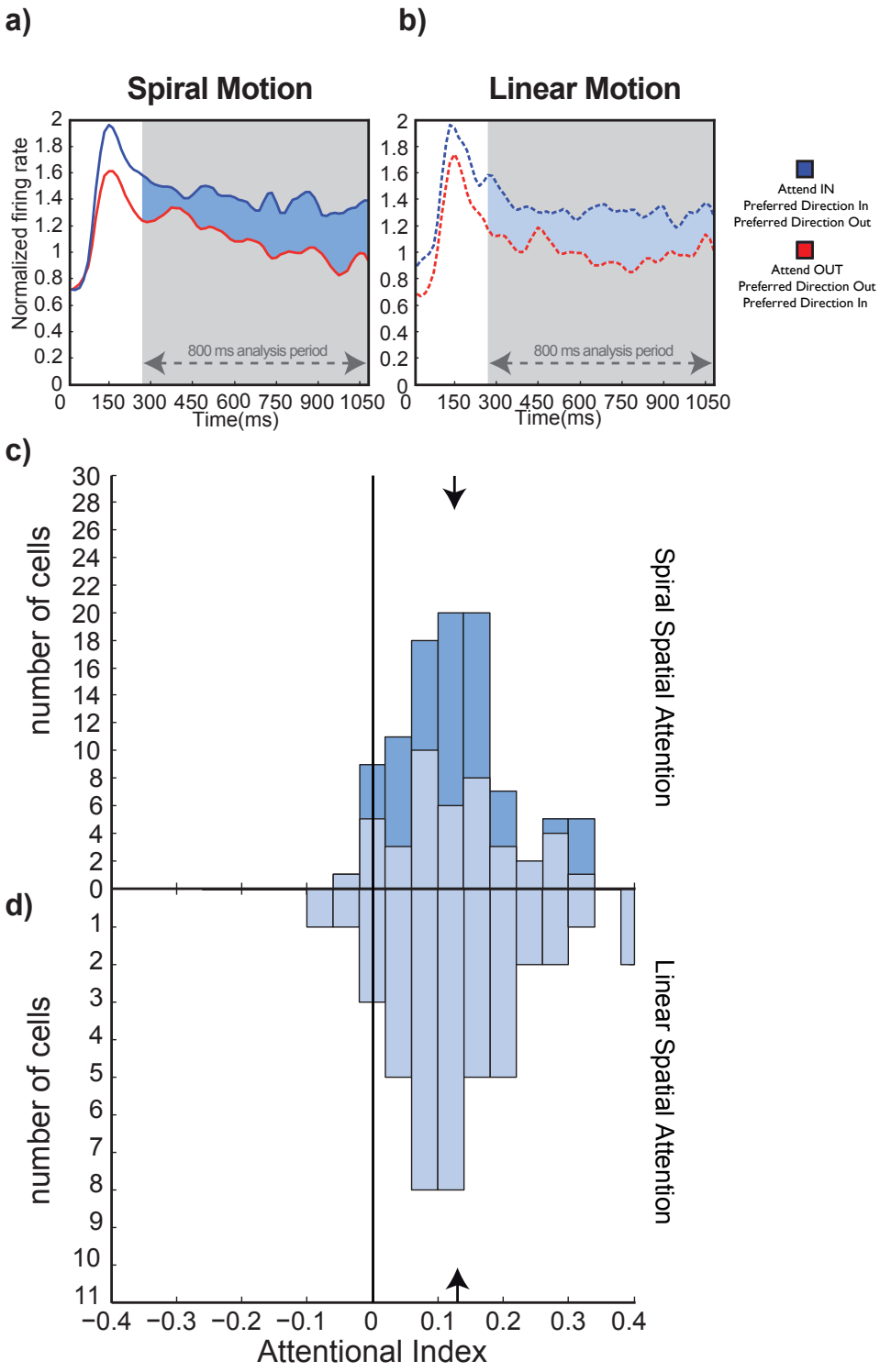


Figure 2: Spiral motion and linear motion spatial attention. The top row shows the time courses of attentional modulation during the spatial attention task; attended preferred motion stimulus (blue trace) versus unattended preferred stimulus (red trace) inside RF. The activity of the recorded population in the SMS (a) (104 neurons) and linear motion (b) (subset 43 neurons) spatial attention task. The bottom histogram ((c) and (d)): distribution of neuronal attentional modulation for MSTd cells. The modulation is expressed as attentional index. Figure (c) MSTd modulation to SMS motion stimuli the

arrow indicates the mean of the distribution (0.13) (the highlighted light blue bars display the subset of 43 neurons recored with both SMS and linear spatial attention). Figure (d) MSTd linear motion spatial attention modulation (mean 0.14 (arrow)).

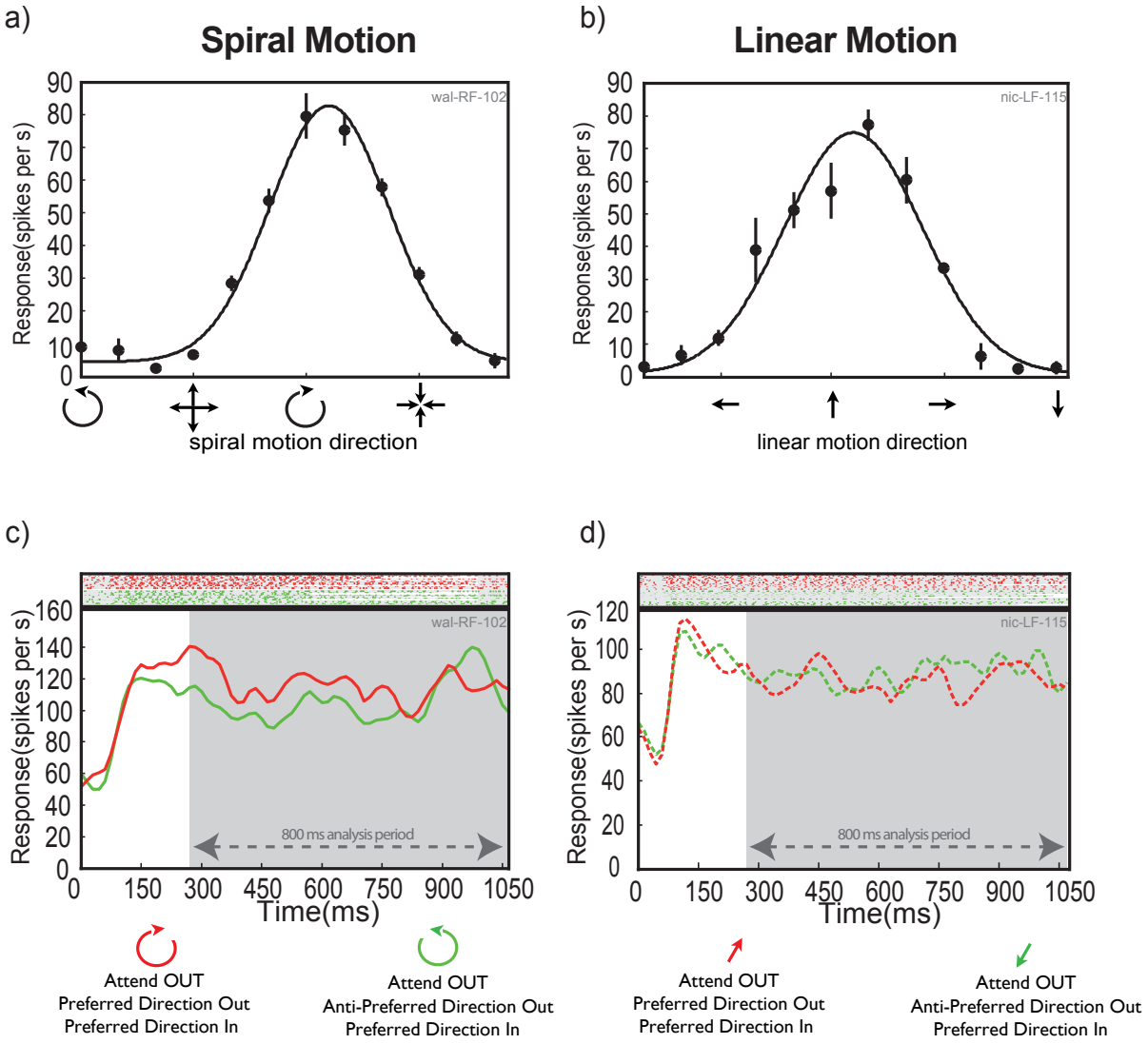


Figure 3: Feature-based attention single cell examples. The top row shows the tuning of two recorded MSTd neurons ((a) SMS tuning; (b) linear motion tuning). Neurons were chosen to have similar firing rates: the preferred SMS direction of cell (a) was a clockwise spiral while cell (b) preferred a linear motion stimulus moving up toward the right. The time courses of the cells feature-based attentional modulations are depicted in the bottom row ((c) SMS feature-based attention and (d) linear motion feature-based attention). A list of spike trains of each trial acquired during a single trial of attending to the preferred direction outside the RF (red) versus attending the anti-preferred direction outside the RF (green) with an unattended motion stimulus moving in the preferred direction placed inside the RF is shown above the convolved, averaged across trials spike density function.

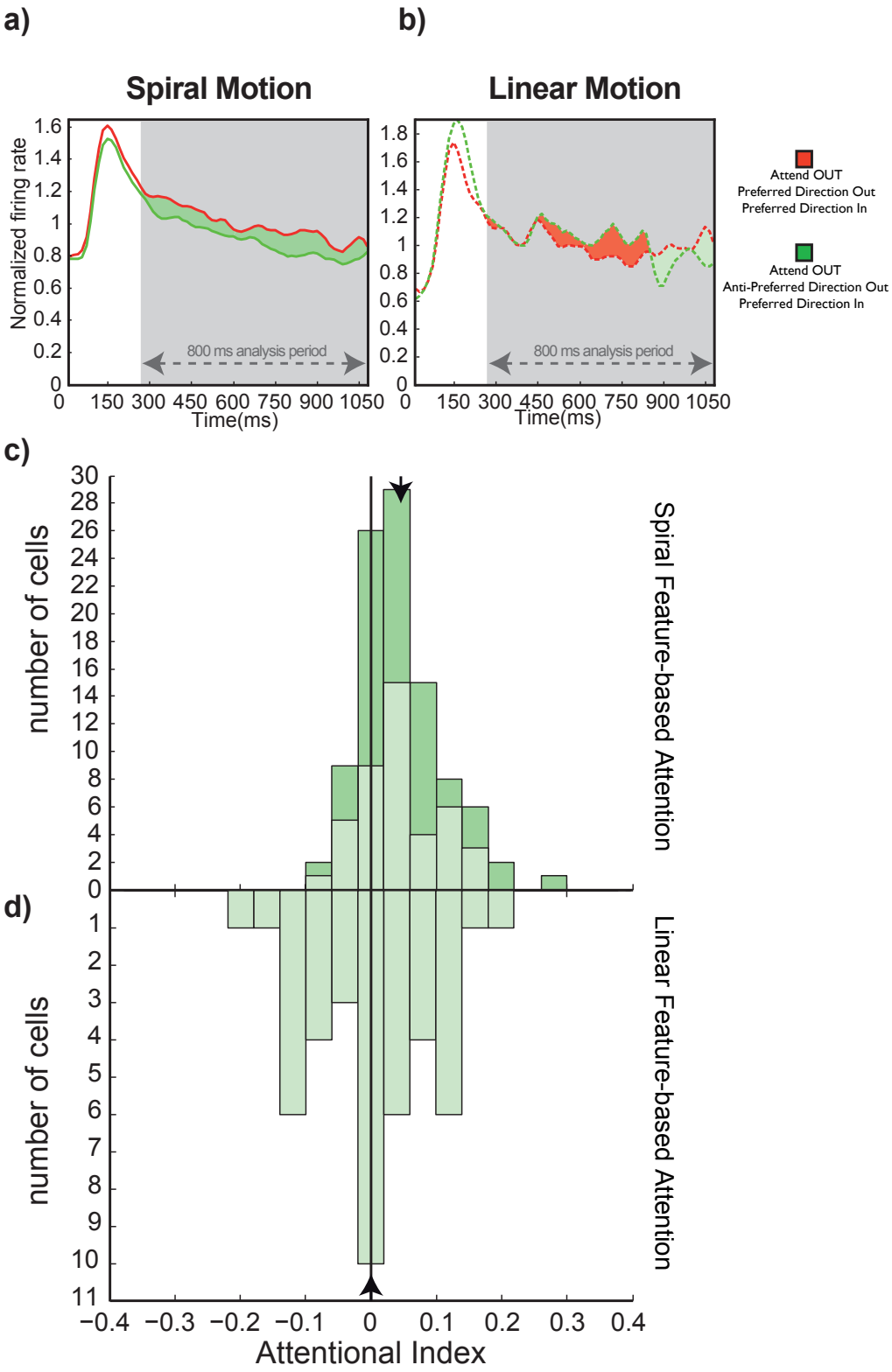


Figure 4: Spiral motion and linear motion feature-based attention. The top row shows the time courses of attentional modulation; attended preferred motion stimulus (red trace) versus attended anti-preferred stimulus (green trace) placed outside the RF. The population responses of feature-based in the SMS (a) (104 neurons) and linear motion (b) (subset 43 neurons) condition. The attentional index histograms ((c) and (d)) show the modulation of MSTd neurons during the feature-based attention task. Attending to the preferred SMS shifted the mean to 0.05 (c) while no significant shift in the AI can be observed for linear tasks. The light green bars (plot c) indicate cells recorded with both SMS and linear feature-based attention.

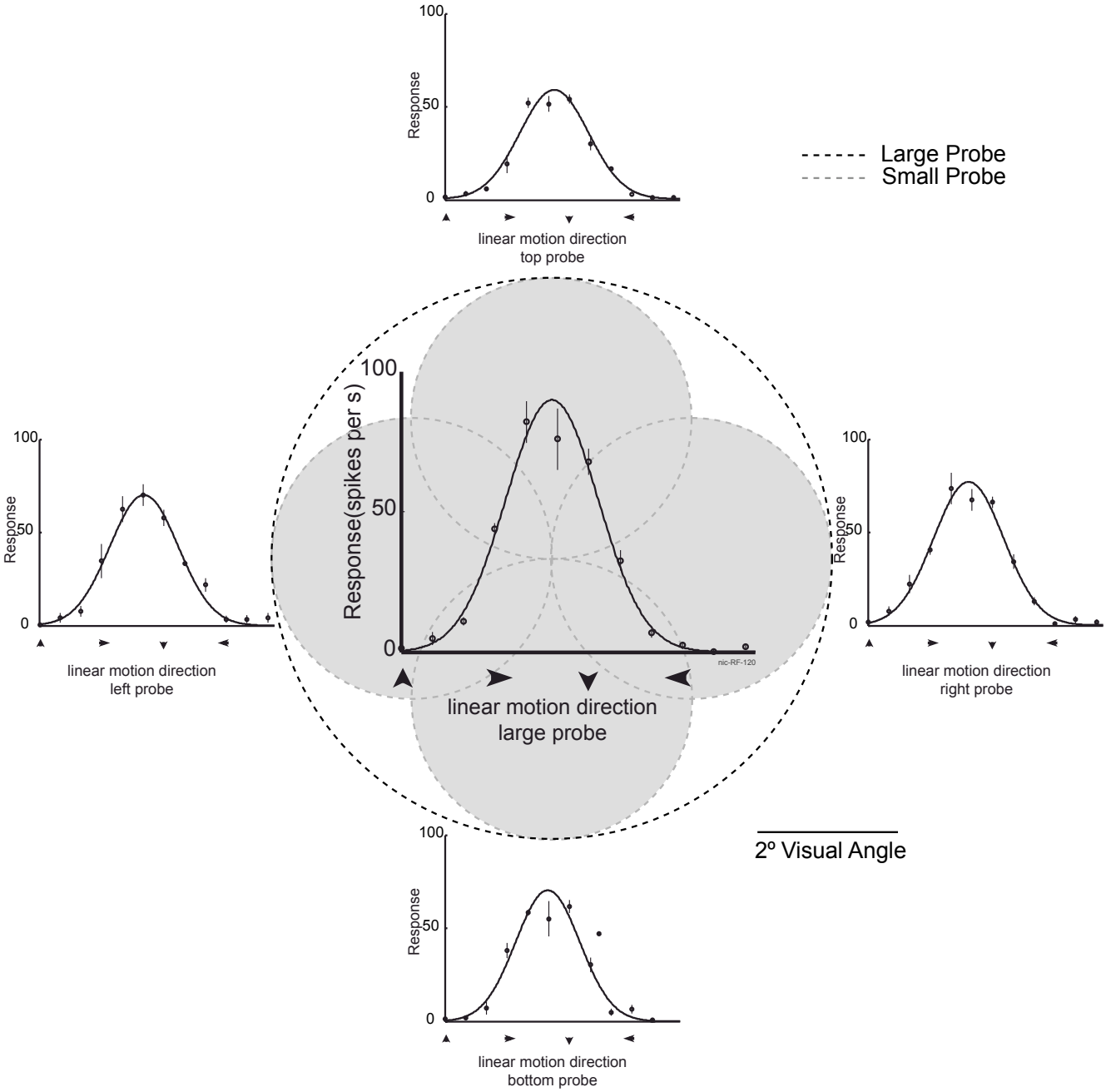


Figure 5: Linear motion position invariance test. At each trial a linear motion stimulus appeared at one of five positions (random presentation). The tuning curve in the center was recorded during the presentation of the large probe stimulus (black dotted circle) matching size and location of the stimuli used during the attentional tasks. Responses of the four small probe stimuli are pictured clockwise starting with the top probe, right probe, bottom probe followed by the left probe. All the stimulus positions are drawn to scale with respect to the large probe stimulus. The small probe stimuli (gray circles) covered half the diameter of the large probe stimulus.

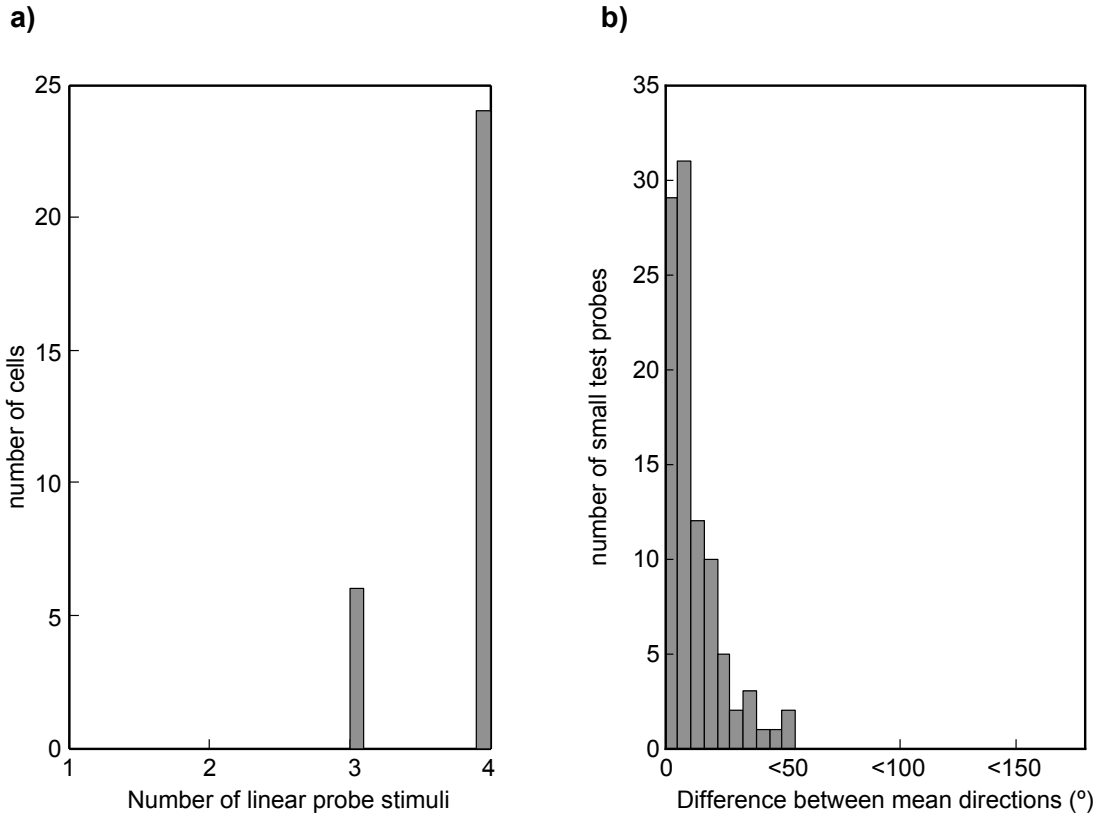


Figure 6: Linear motion position invariance test. Bar graph (a) shows the number of probe stimuli for each cell with a direction selectivity of three times higher for the preferred direction over the anti-preferred direction. 24 cells have all four small probe stimuli responding three times higher. Bar graph (b) plots the absolute difference between the small probe linear motion stimuli with respect to the large reference probe stimulus.

References

Brefczynski JA, DeYoe EA (1999). A physiological correlate of the 'spotlight' of visual attention. Nat Neurosci; 2 (4): 370-4.

Cameron EL, Tai JC, Carrasco M (2002). Covert attention affects the psychometric function of contrast sensitivity. Vision Res;42(8): 949-967.

Corbetta M, Miezin FM, Shulman GL, Petersen SE (1993). A PET study of visuospatial attention. J Neurosci; 13 (3): 1202-26.

Connor CE, Preddie DC, Gallant JL, Van Essen DC (1997). Spatial attention effects in macaque area V4. J Neurosci; 17 (9): 3201-14.

Cook EP, Maunsell JH (2002). Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. J Neurosci; 22 (5): 1994-2004.

Desimone R, Duncan J(1995). Neural mechanisms of selective visual attention. Annu Rev Neurosci; 18: 193-222.

Duffy CJ, Wurtz RH (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J Neurophysiol; 65 (6): 1329-45.

Eriksen BA, Eriksen CW (1974). Effect of noise letters upon the identification of a target letter in a nonserach task. Percept Psychophys; 16(1): 143-149.

Eriksen CW, Hoffman, JE (1973). The extent of processing of noise elements during selective encoding from visual display. Percept Psychophys; 14: 155-160.

Felleman DJ, Van Essen DC (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex; 1 (1): 1-47

Gandhi SP, Heeger DJ, Boynton GM (1999). Spatial attention affects brain activity in human primary visual cortex. Proc Natl Acad Sci U S A; 96 (6): 3314-9.

Geesaman BJ, Andersen RA (1996). The analysis of complex motion patterns by form/ cue invariant MSTd neurons. J Neurosci; 16 (15): 4716-32.

Graziano MS, Andersen RA, Snowden RJ (1994). Tuning of MST neurons to spiral motions. J Neurosci; 14 (1): 54-67.

Hurlbert A, Poggio T (1985). Spotlight on attention. Trends Neurosci.;8:309 –311.

Katzner S, Busse L, Treue S (2009). Attention to the Color of a Moving Stimulus Modulates Motion-Signal Processing in Macaque Area MT: Evidence for a Unified Attentional System. Front Syst Neurosci; 3: 12.

Lu ZL, Dosher B A(1998). External noise distinguishes attention mechanisms. Vision Res; 38(9): 1183-1198.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol; 77 (1): 24-42.

McAdams CJ, Maunsell JH (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci; 19 (1): 431-41.

McAdams CJ, Maunsell JH (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. J Neurophysiol; 83 (3): 1751-5.

McAdams CJ, Reid RC (2005). Attention modulates the responses of simple cells in monkey primary visual cortex. J Neurosci; 25 (47): 11023-33.

McAlonan K, Cavanaugh J, Wurtz RH (2008). Guarding the gateway to cortex with attention in visual thalamus. Nature; 456 (7220): 391-4.

Maunsell, JHR (2004). The role of attention in visual cerebral cortex. (In) L.M. Chalupa and J.S. Werner (Eds.) The Visual Neurosciences. MIT Press, Cambridge MA.: 1538-1545.

Moran J, Desimone R (1985). Selective attention gates visual processing in the extrastriate cortex. Science; 229 (4715): 782-4.

Motter BC (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol; 70 (3): 909-19.

Motter BC (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. J Neurosci; 14 (4): 2178-89.

Orban GA, Lagae L, Verri A, Raiguel S, Xiao D, Maes H, Torre V (1992). First-order analysis of optical flow in monkey brain. Proc Natl Acad Sci U S A; 89 (7): 2595-9.

Posner MI, Snyder CR, Davidson BJ (1980). Attention and the detection of signals. J Exp Psychol; 109 (2): 160-74.

Saito H, Yukie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J Neurosci; 6 (1): 145-57.

Shiu LP, Pashler H (1995). Spatial attention and vernier acuity. Vision Res; 35(3): 337-343.

Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. J Neurosci; 6 (1): 134-44.

Treue S (2003a). Climbing the cortical ladder from sensation to perception. Trends Cogn Sci; 7 (11): 469-71.

Treue S (2003b). Visual attention: the where, what, how and why of saliency. Curr Opin Neurobiol; 13 (4): 428-32.

Treue S, Martínez Trujillo JC (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature; 399 (6736): 575-9.

Treue S, Maunsell JH (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. Nature; 382 (6591): 539-41.

Tootell RB, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM (1998). The retinotopy of visual spatial attention. Neuron; 21 (6): 1409-22.

Ungerleider LG, Desimone R (1986). Cortical connections of visual area MT in the macaque. J Comp Neurol.;248 (2): 190-222

Yeshurun Y, Carrasco M (1998). Attention improves or impairs visual performance by enhancing spatial resolution. Nature; 396(6706): 72-75.

Chapter IV

Summary

In summary, this thesis has investigated the influence of adaptation and attention on “higher” visual perception. While it is clear that the visual cortex includes functionally defined regions that can be hierarchically ordered (Felleman and Van Essen, 1991) on the basis of information flow and processing, the term “higher” visual perception is used loosely. Throughout this thesis it refers to the perceptual processing of objects, scenes and faces as well as cortical regions beyond the occipital lobe in which retinotopical organization is generally not conserved (e.g. FFA fusiform face area, MSTd dorsal region of the medial superior temporal).

The first part of this work focuses on adaptation, a basic phenomenon related to sensory processing. Psychophysically tested, showing a particular stimulus such as an artificially generated image mimicking the statistical characteristics of a (i) natural or (ii) man-made environment leads to a perceptual shift / recalibration: neutral (ambiguous) rapidly presented test stimuli appear to have opposite characteristics ((i) man-made or (ii) natural scene). We suggest that the classification of rapidly presented man-made and natural images can routinely be influenced by the statistical properties of the prevailing scene. Consequently the categorization of “real world” environments can be carried out nearly effortlessly employing a simple feedforward mechanism making use of very basic information. A physiological prediction is that spatial frequency adaptation along different orientations will not only depress or otherwise affect responses of V1 cells but also responses in “high-level” areas encoding complex visual scenes. In a complementary experiment we investigated how adaptation adjusts the sensitivity of rapid parallel scene perception and whether the categorization of natural vs. man-made scenes reflects distinct processing boundaries. We provide evidence that category information can rapidly be extracted as a result of removing redundant information. This could give rise to a stimulus-related pre-recognition ahead of complex object based scene recognition. Further, our results reveal distinct processing boundaries contending

opponent processing of natural and man-made scenes. Finally we examined the neural correlates of face adaptation by monitoring BOLD responses with fMRI. Asymmetric response changes in face selective areas, with higher BOLD signals when a normal face followed a distorted face than vice versa parallels the relative salience of the perceptual face distortion aftereffect. Our results suggest that the aftereffects for face adaptation reflect a recentering of the neural representations for faces relative to a set norm.

A second focus of this thesis are the attentional influences on the processing of complex motion patterns. Using extracellular single-unit recordings in motion processing area MSTd
of awake behaving macaque monkeys, we show that spatial attention modulates the response of MSTd neurons to optimized spiral motion stimuli in a similar multiplicative fashion as previously reported for other extrastriate areas (e.g. MT - Treue and Martínez Trujillo, 1999; V4 McAdams and Maunsell, 1999). Finally, results of the most recent electrophysiological investigation of feature-based attention in area MSTd are presented. Here we show that feature-based attention multiplicatively modulates the response of MSTd neurons to complex spiral motion by the same factor as the responses of MT neurons to linear motion. While MSTd neurons respond well to linear motion no feature-based attentional modulation for linear motion stimuli can be reported.
Together, these contributions show that both adaptation and attention strongly influence visual processing.

Bibliography

Albright TD, Desimone R and Gross CG (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. Journal of Neurophysiology, 51 (1): 16-31

Anton-Erxleben K, Stephan VM and Treue S (2009). Attention reshapes center-surround receptive field structure in macaque cortical area MT. Cereb Cortex.;19(10): 2466-78

Barbas H and Mesulam MM(1981) Organization of affer- ent input to subdivisions of area 8 in the rhesus monkey. J. Comp. Neurol. 200: 407-431.

Barlow, H B (1990). A theory about the functional role and synaptic mechanisms of visual aftereffects. Vision coding and efficiency Edited Blake, C Cambridge University Press, UK

Boussaoud D, Desimone R and Ungerleider LG (1991). Visual topography of area TEO in the macaque. J Comp Neurol.;306(4):554-75

Boussaoud D, Ungerleider LG and Desimone R (1990). Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. J Comp Neurol.;296 (3): 462-95

Britten KH, Shadlen MN, Newsome WT and Movshon JA (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci.; 12: 4745–4765

Carandini, M (2000). Visual cortex: Fatigue and adaptatio. Current Biology, 10 (16):R605–R607

Clifford CWG, Webster MA, Stanley GB, Stocker AA, Kohn A, Sharpee TO and Schwartz O (2007). Visual adaptation: Neural, psychological and computational aspects. Vision Research, 47: 3125–3131

Colby CE and Goldberg ME (1999). Space and attention in parietal cortex. Annual Review of Neuroscience, 22, 319–349

DeAngelis GC and Newsome WT (1999). Organization of Disparity-Selective Neurons in Macaque Area MT. J. Neurosci.;19 (4): 1398-1415

Desimone R and Ungerleider LG (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. J Comp Neurol.; 248 (2):164-89

Duffy CJ and Wurtz RH (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. Journal of Neurophysiology; 65(6):1329-45

Dursteler and Wurtz RH (1988). Pursuit and Optokinetic Deficits Following Chemical Lesions of Cortical Areas MT and MS. Journal of Neurophysiology;60 (3): 940-965

Exner S (1894). Entwurf zu einer physiologischen Erklärung der psychischen Erscheinung (Ostwalds Klassiker der exakten Wissenschaften; Bd. 285). Frankfurt am Main, Germany, 1999

Fries P (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci.;9(10):474-48017

Fries P, Reynolds JH, Rorie AE and Desimone R (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. Science.;291(5508):1506-7.

Fries W, Keizer K and Kuypers HGJM (1985). Layer VI cells in the macaque striate cortex project to both superior colliculus and orestriate visual area V5. EXD. Brain Res. 58: 6 13-6

Galletti C, Battaglini PP and Fattori P (1995). Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. The Euro. J Neurosci.;7(12):2486-501

Geesaman BJ, Andersen RA (1996). The analysis of complex motion patterns by form/cue invariant MSTd neurons. J Neurosci.; 16 (15): 4716-32

Graham N (1972).Spatial frequency channels in the human visual system: effects of luminance and pattern drift rate. Vision Res. 12, 53–68

Graziano MS, Andersen RA and Snowden RJ (1994). Tuning of MST neurons to spiral motions. J Neurosci.;14 (1):54-67

Goldberg, M. E. and R. H. Wurtz (1972). "Activity of the superior colliculus in behaving monkey II. Effect of attention on neuronal responses." JNP 35: 560-574.

Hawken MJ, Parker AJ, and Lund JS (1988). Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the Old World monkey. J Neurosci.; 8: 3541–3548

Heuer HW and Britten KH (2004). Optic flow signals in extrastriate area MST: comparison of perceptual and neuronal sensitivity. J Neurophysiol.;91(3):1314-26

Hubel DH and Wiesel TN (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. London 195(1) pages 215-243

Hubel DH and Wiesel TN (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. J Comp Neurol.;146(4):421-50

Ilg UJ, Schumann S, Thier P(2004). Posterior parietal cortex neurons encode object motion in world-centered coordinates. Neuron; 43: 145-151

James W (1980). The principles of psychology. Dover Publications [1950, c1918] NewYork, New York USA

Komatsu, H. and R. H. Wurtz (1989). "Modulation of pursuit eye movements by stimulation of cortical areas MT and MST." JNP 62(1): 31-47

Komatsu, H. and R. H. Wurtz (1988). "Relation of cortical areas MT and MST to pursuit eye movements I. Localization and visual properties of neurons." JNP 60(2): 580-603

Li CY, Tanaka M and Creutzfeldt OD (1989). Attention and eye movement related activation of neurons in the dorsal prelunate gyrus (area DP). Brain Res. 1989 Sep 4;496(1-2):307-13

Luck SJ, Chelazzi L, Hillyard SA and Desimone R (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol.; 77(1):24-42

Lund JS, Lund RD, Hendrickson AE, Bunt AH and Fuchs AF (1976). The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by retrograde transport of horseradish peroxidase. J. Comp. Neurol. 164: 287-304

McAdams CJ, Maunsell JH (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci.;19(1):431-41

Maunsell, JHR (2003). The role of attention in visual cerebral cortex. VN. L. W. Chalupa and J. S. Werner. Cambridge MA, MIT Press: 1538-1545

Maunsell, JHR and McAdams CJ (2001). Effects of attention on the responsiveness and selectivity of individual neurons in visual cerebral cortex. Visual attention and cortical circuits. J. Braun, C. Koch and J. L. Davis. Cambridge, Massachusetts, MIT Press: 103-120

Maunsell JHR and Van Essen DC(1983). The connections of the middle temporal visual area (MT) and their relationship to cortical hierarchy in the macaque monkey. J. Neurosci. 3: 2563-2586.

Maunsell, JHR and Treue S(2006). "Feature-based attention in visual cortex." TINS 29(6): 317-322

Mountcastle, VB (1976). "Then world around us: Neural command functions for selective attention." Neurosci. Res. Prog. Bull. 14: 1-47

Movshon JA, Lennie P (1979). Pattern-selective adaptation in visual cortical neurons. Nature, 278:850-852

Newsome, W. T., Wurtz R. H. and Komatsu, H. (1988). "Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs." JNP 60(2): 604-620

Orban GA, VanEssen D and Vanduffel W (2004). Comparative mapping of higher visual areas in monkeys and humans. Trends Cog. Sci.;.8(7): 315-324

Petersen S E, Baker J F and Allman J M (1985). Direction-specific adaptation in area MT of the owl monkey. Brain Research, 346 (1):146-150

Posner MI, Snyder CR and Davidson BJ (1980). Attention and the detection of signals. J Exp Psychol.;109(2):160-74

Raffi M and Siegel RM (2007). A functional architecture of optic flow in the inferior parietal lobule of the behaving monkey. PLoS ONE;2 (2)e200:1-19

Rhodes G, Jeffery L, Watson TL, Clifford CW and Nakayama K (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. Psychol Sci.;14(6):558-66

Rosenberg A, Wallisch P and Bradley DC (2008). Responses to direction and transparent motion stimuli in area FST of the macaque. Visual Neurosci.; 25(2):187-195

Saito HA, Yukie M, Tanaka K, Hikosaka K, Fukada Y and Iwai E (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J Neurosci.;6 (1):145-57

Seidemann E and Newsome WT (1999). Effect of spatial attention on the responses of area MT neurons. J Neurophysiol.;81(4):1783-94

Sekuler R and Pantle A (1967). A model for after-effects of seen movement. Vision Research, 7:427-439

Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. 18, 3870–3896

Spitzer H, Desimone R and Moran J (1988).Increased attention enhances both behavioral and neuronal performance. Science.;240(4850):338-40

Tammer R, Hofer S, Merboldt KD, Frahm J. Magnetic Resonance Imaging of the Rhesus Monkey Brain. Vandenhoeck & Ruprecht, Göttingen, 1. Ed. 2009.

Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. J Neurosci.;6 (1):134-44

Tanaka K, Saito HA (1989). Analysis of motion of the visual field by direction, expansion/ contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. Journal of Neurophysiology; 62 (3):626-41

Tootell R B, Reppas J B, Dale A M, Look R B, Sereno M I, Malach R, Brady TJ , Rosen B R (1996). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. Nature;375(6527):139-41

Treue, S (2001). "Neural correlates of attention in primate visual cortex." TINS 24(5): 295-300

Treue, S (2003). "Visual attention: the where, what, how and why of saliency." CONB 13: 428-432

Treue, S and JC Martinez-Trujillo (1999). "Feature-based attention influences motion processing gain in macaque visual cortex." Nature 399(6736): 575-579

Treue S and Maunsell JH (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. Nature.;382(6591):539-41

Ungerleider LG and Desimone R (1986). Cortical connections of visual area MT in the macaque. J Comp Neurol.;248 (2): 190-222

Van Essen DC, Anderson CH and Felleman DJ(1992). Information processing in the primate visual system: an integrated systems perspective.Science;255(5043):419-23

Van Essen DC, Maunsell JHR and Bixby JL(1981). The middle temporal visual area in the macaque: Myeloarchitecture, connections, functional properties and topographic organization. J. Comp. Neurol.; 199: 293-326.

Van Essen DC, Harwell J, Hanlon D and Dickson J (2005). Surface-Based Atlases and a Database of Cortical Structure and Function. In: Databasing the Brain: From Data to Knowledge (Neuroinformatics). S.H. Koslow and S. Subramaniam, eds., John Wiley & Sons, NJ, pp. 369-387

Vautin R G and Berkley M A (1977). Responses of Single Cells in Cat Visual Cortex to Prolonged Stimulus Movement: Neural Correlates of Visual Aftereffects. Journal of Neurophysiology, 40 (5): 1051-1065

Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. Cereb Cortex.;3(2): 79-94

Watson TL and Clifford CW (2003). Pulling faces: an investigation of the face-distortion aftereffect. Perception;32(9):1109-16

Webster M A, and McLin OH (1999). Figural aftereffects in the perception of faces. Psychonomic Bulletin & Review, 6, 647-653

Webster M A and Mollon J D (1994). The influence of contrast adaptation on color appearance. Vision Res.;34(15):1993-2020

Webster M A, Werner J S, and Field D J (2005). Adaptation and the phenomenology of perception. In Fitting the Mind to the World: Adaptation and Aftereffects in High Level Vision, Clifford and Rhodes, eds. (Oxford: Oxford University Press)

Williford T and Maunsell JHR (2006). Effects of spatial attention on contrast response functions in macaque area V4. Journal of Neurophysiology, 96: 40-54

Womelsdorf T, Anton-Erxleben K, Pieper F and Treue S (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. Nat Neurosci.;9(9):1156-60

Womelsdorf, T. and Fries P (2007). "The role of neuronal synchronization in selective attention." CONB 17: 154-160

Wurtz RH, Goldberg ME, Robinson DL (1980). Behavioral modulation of visual responses in the monkey: Stimulus selection for attention and movement. Progr. Psychobio. and Phys. Psych.;9: 43-83.

Yin, T. C. T. and Mountcastle V B(1978). "Mechanisms of neural integration in the parietal lobe for visual attention." FP 37: 2251-2257

Zeki SM (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J. Physiol.; 236: 549-573

Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS (1991). A direct demonstration of functional specialization in human visual cortex. J Neurosci.;11(3): 641-9

Curriculum Vitae

Daniel Kaping

German Primate Center
Cognitive Neuroscience Laboratory

Kellnerweg 4
37077 Göttingen
Germany
Phone: +49 (551) 3851-354
Fax:+49 (551) 3851-183
E-mail: dkaping@gwdg.de

Education

since 2004 **PhD student - Center for Systems Neuroscience, Graduate School for Neuroscience and Molecular Biosceinces, Göttingen, Germany,** Ph.D. Student, working under Professor Stefan Treue at the German Primate Center, Cognitive Neuroscience Laboratory

2001-2004 **University of Nevada, Reno,** M.A. Experimental Psychology (Cognitive, Developmental/Comparative, and Brain Science); working under Professor Michael A. Webster
GPA: 3.81

1998-2001 **University of Nevada, Reno,** B.A. Psychology, GPA: 3.4 in Major: 3.6. Minor: Criminal Justice

Educational Experiences

2008 **Advanced Course in Computational Neuroscience**
Bernstein Center for Computational Neuroscience
Freiburg, Germany August 3 – 28. 2008

2008 **Experimental Animal Course**
Georg-August-University Göttingen, Medicine
Göttingen, Germany October 6 – 10. 2008

2004 – 2008 **SensoPrim** (sensory information processing in non-human primates)
Marie Curie Fellowship - Early Stage Research Training Program (EST)

February 2007 - "Sensory-motor systems and neuroprosthetics",
Allg. Zoologie und Neurobiologie,
Ruhr-Universität Bochum, Germany

June 2006 - “Communication Training”,
Department of Neurology,
Universitäts-Klinikum
Munich, Germany

January 2006 - “The making of science”, Department of Biophysics, Radboud University of Nijmegen,
The Netherlands

September 2005 - “Perspectives of Primatology”,
German Primate Center,
Göttingen, Germany

July 2005 - “Introduction to MATLAB”,
Department of Neurology,
Universitäts-Klinikum
Munich, Germany

2004 **Course on Computational Neuroscience**
Max Planck Institute for Dynamics and Self-Organization
Göttingen, Germany September 22 – 26. 2004

2003 **Research Internship**
Max Planck Institute for Biological Cybernetics
Cognitive and Computational Psychophysics
Prof. Dr. Heinrich Bülthoff
Object recognition & face perception

2001-2004 **Research Assistant**
University of Nevada, Reno
Cognitive, Developmental/Comparative, and Brain Science
Perception Lab (Michael Webster)
Face perception and the effects of adaptation

Teaching Experience

Spring 2009 **Neuro-Biology**
University of Göttingen, Germany

Spring 2008 **Neuro-Biology**
University of Göttingen, Germany

Summer 2004 **General Psychology**
University of Nevada, Reno, USA

Summer 2003 **Statistics for Social Science**
University of Nevada, Reno, USA

Publications

Kaping, D., Tzvetanov, T., Treue, S. (2007). *Adaptation to statistical properties of visual scene biases rapid categorization*. Visual Cognition 15(1), 12-19

Graf, M., Kaping, D., Bülthoff, H.H. (2005). *Orientation Congruency Effects for Familiar Objects. Coordinate Transformation in Object Recognition*. Psychological Science 16(3), 214-222

Webster, M.A., Kaping, D., Mizokami, Y., Duhamel, P. (2004). *Adaptation to Natural Facial Categories*. Nature 428, 557-561

Conference Abstracts

Dechent P., Morawetz C., Baudewig J., Treue S., Webster M. A., Kaping D. (2009) Faces around the norm, fMRI of the Face Distortion Aftereffect. 17th Scientific meeting International Society for Magnetic Resonance in Medicine, Honolulu, Hawaii, USA, May 2009

Kaping D., Baloni, S, Treue S. (2009) Spatial attention in area MST. 2nd Annual Primate Neurobiology Meeting, Göttingen, Germany, March 2009

Kaping D., Baloni, S., Treue S. (2008) Multiplicative attentional modulation of MSTd tuning to spiral motion patterns. Society for Neuroscience, Washington D.C., USA, November 2008

Kaping D., Treue S. (2008) fMRI face distortion aftereffect revealing norm based face coding. BCCN Excellence Workshop "Aspects of Adaptive Cortical Dynamics", Delmenhorst, Germany, September 2008

Kaping D., Treue S. (2008) Task-specific attentional modulation of MT responses. 1st Annual Primate Neurobiology Meeting, Tübingen, Germany, February 2008

Kaping D., Tchumatchenko T., Geisel T., Treue S. (2007) Task-specific attentional modulation of MT resposnes. Society for Neuroscience, San Diego, California, USA, November 2007

Morawetz C., Kaping D., Baudewig J., Treue S., Webster M. A., Dechent P. (2007) Neural mechanisms underlying the face distortion aftereffect. 13th Annual meeting Human Brain Mapping, Chicago, Illinois, USA, June 2007

Kaping D., Morawetz C., Baudewig J., Treue S., Webster M. A., Dechent P. (2007) Face distortion aftereffect activates motion and face sensitive areas: an fMRI study. Vision Sciences Society, Sarasota, Florida, USA, May 2007

Kaping, D., Tzvetanov, T., Treue, S. (2007) Categorization of visual scenes based on low-level image statics. 7th Meeting of the German Neuroscience Society/31st Goettingen Neurobiology Conference, Goettingen, Germany, April 2007

Kaping, D., Tzvetanov, T., Treue, S. (2006) Effect of adaptation suggests role of low-level processes in rapid scene categorization. Vision Sciences Society, Sarasota, Florida, USA, May 2006 – Journal of Vision, 6 (6), Abstract 800

Tzvetanov, T. & Kaping, D. (2005) A more precise parameters estimation of short length Markov chain data sets. 7 Colloqueede la Société des Neurosciences, Lille, France, May 2005

Yasuda, M., Bedard, K., Mizokami, Y., Kaping, D., Webster M.A. (2005) Adaptation and individual differences in categorical judgments of faces. Vision Sciences Society, Sarasota, Florida, USA May 2005 - Journal of Vision, 5 (8), Abstract 832

Paras. C., Kaping, D., Webster M.A. (2004) Adaptation and the perception of facial symmetry. Vision Sciences Society, Sarasota, Florida, USA May 2004 – Journal of Vision, 4 (8), Abstract 440

Graf, M., Kaping, D., Bülthoff H.H. (2003) Reference Frame Adjustment In Object Recognition. Conference of the Psychonomic Society. Vancouver, Canada November 2003

Graf, M., Kaping, D., Bülthoff, H.H. (2003) Orientation congruency effect in object recognition. European Conference on Visual Perception, Paris, France September 2003

Ng, M., Kaping, D., Anstis, S., Fine, I., Webster, M.A. (2003) Selective tuning of face perception. Vision Sciences Society, Sarasota, Florida, USA May 2003 – Journal of Vision, 3 (9), Abstract 106

Kaping, D., Mizokami, Y., Webster, M.A. (2003) Adapting to a new visual environment: A field study of face perception. Vision Sciences Society, Sarasota, Florida, USA May 2003 - Journal of Vision, 3 (9), Abstract 296

Kaping, D., Duhamel, P., Webster, M.A. (2002) Adaptation to Natural Facial Categories. Fall Vision Meeting, San Francisco, California, USA Oct 2002 – Journal of Vision, 2 (10), Abstract 128

Kaping, D., Bilson, A.C., Webster, M.A. (2002) Adaptation and Categorical Judgments of Faces. Vision Sciences Society, Sarasota, Florida, USA May 2002 - Journal of Vision, 2 (7), Abstract 564

Bilson, A.C., Kaping, D., Webster, M.A. (2002) Stimulus Configuration Supporting the Perception of Faces. Vision Sciences Society, Sarasota, Florida, USA May 2002 - Journal of Vision, 2 (7), Abstract 621

Kaping, D., Bilson, A.C., Webster, M.A. (2002) Adaptation Effects on Natural Categorical Judgments of Faces. Rocky Mountain Psychological Association, Park City, Utah, USA _

Göttingen, 31.12. 2009